

Ab-initio multimode linewidth theory for arbitrary inhomogeneous laser cavities

A. Pick,¹ A. Cerjan,² D. Liu,³ A. W. Rodriguez,⁴ A. D. Stone,² Y. D. Chong,⁵ and S. G. Johnson⁶

¹*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*

²*Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA*

³*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

⁴*Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA*

⁵*Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore*

⁶*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

We present a multimode laser-linewidth theory for arbitrary cavity structures and geometries that contains nearly all previously known effects and also finds new nonlinear and multimode corrections, e.g. a correction to the α factor due to openness of the cavity and a multimode Schawlow–Townes relation (each linewidth is proportional to a sum of inverse powers of all lasing modes). Our theory produces a quantitatively accurate formula for the linewidth, with no free parameters, including the full spatial degrees of freedom of the system. Starting with the Maxwell–Bloch equations, we handle quantum and thermal noise by introducing random currents whose correlations are given by the fluctuation–dissipation theorem. We derive coupled-mode equations for the lasing-mode amplitudes and obtain a formula for the linewidths in terms of simple integrals over the steady-state lasing modes.

I. INTRODUCTION

The fundamental limit on the linewidth of a laser is a foundational question in laser theory [1–5]. It arises from quantum and thermal fluctuations [6, 7], and depends on many parameters of the laser (materials, geometry, losses, pumping, etc.); it remains an open problem to obtain a fully general linewidth theory. In this paper, we present a multimode laser-linewidth theory for arbitrary cavity structures and geometries that contains nearly all previously known effects [8–12] and also finds new nonlinear and multimode corrections. The theory is quantitative and makes no significant approximations; it simplifies, in the appropriate limits, to the Schawlow–Townes formula (2) with the well-known corrections. It also demonstrates the interconnected behavior of these corrections [13, 14], which are usually treated as independent. Most previous laser-linewidth theories have employed simple models for calculating the lasing modes (e.g., making the paraxial approximation). Such simplifications, though appropriate for many macroscopic lasers, are inadequate for describing complex microcavity lasers such as 3d nanophotonic structures or random lasers with inhomogeneities on the wavelength scale [15–18]. We base our theory on the recent steady-state ab-initio laser theory (SALT) [19, 20], which allows us to efficiently solve the semi-classical laser equations in the absence of noise for arbitrary structures [21]. We treat the noise as a small perturbation to the SALT solutions, allowing us to obtain the linewidths *analytically* in terms of simple integrals over the steady-state lasing modes. Our SALT-based theory is *ab initio* in the sense that it produces quantitatively accurate formulas for the linewidths, with no free parameters, including the full spatial degrees of freedom of the system. Hence, we will refer to this approach as the noisy steady-state *ab-initio* laser theory (N-SALT).

Our derivation (Secs. III–V) begins with the Maxwell–Bloch equations (details in appendix A), which couple the full-vector Maxwell equations to an atomic gain medium [22], combined with random currents (in Sec. IV) whose statistics are described by the fluctuation–dissipation theorem (FDT) [23–27]. In the presence of these random currents, the amplitudes of the lasing modes evolve according to a set of coupled ordinary differential equations (ODEs), which have been called “oscillator models” [28, 29] or “temporal coupled-mode theory” (TCMT) [30–34] in similar contexts. In their most general form, our N-SALT TCMT equations (Sec. III) have the form of oscillator equations with a *non-instantaneous* nonlinear term that stabilizes the mode amplitudes around their steady-state values. The non-instantaneous nonlinearity arises since the atomic populations respond with a time delay to field fluctuations; this corresponds to the typical case of “class B” lasers [35–37], in which the population dynamics cannot be adiabatically eliminated. We are able to show analytically that the resulting linewidths of the lasing peaks are identical to the results one obtains for a simplified model with instantaneous nonlinearity [28, 29], which describes the (less common) case of “class A” lasers, in which the population dynamics are adiabatically eliminated. As expected, however, in certain parameter regimes the full non-instantaneous model can exhibit side peaks alongside the main lasing peaks [38], arising from relaxation oscillations (Sec. V.C).

By solving the N-SALT TCMT equations, we obtain a simple closed-form matrix expression for the linewidths and multimode phase correlations (Sec. V), generalizing earlier two-mode results that used phenomenological models [39]. This gives a multimode “Schawlow–Townes” relation (Sec. VI.C), where the linewidth of each lasing mode is proportional to a sum of inverse output powers of the neighboring lasing modes. The theory is valid well

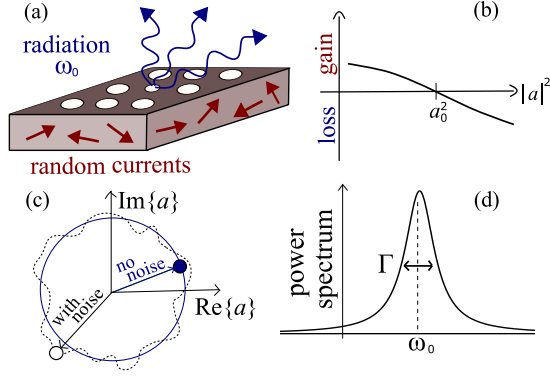


FIG. 1. (Color online) Schematics illustrating linewidth physics. (a) Photonic-crystal (PhC) laser cavity [33] emitting radiation from the lasing mode at frequency ω_0 , perturbed by random currents. (b) The squared amplitude is stabilized around a_0^2 . Below (above) a_0^2 , the medium provides light amplification (attenuation). (c) Phasor diagram for the complex field amplitude: a circular oscillation (with $|a| = a_0$) for the noise-free mode and a perturbed path for noise-driven mode. Noise drives small amplitude fluctuations and possibly large phase drifts. (d) The lineshape is a Lorentzian $\sim \Gamma/[(\omega - \omega_0)^2 + (\frac{1}{2}\Gamma)^2]$, centered around ω_0 with width Γ .

above threshold, and whenever a new mode turns on, this inverse-power relation produces a divergence due to the failure of the linearization approximation near threshold. However, we show that this divergence is spurious and can be avoided by solving the nonlinear N-SALT TCM equations numerically [40]. (Our formalism can be extended to treat the near-threshold regime analytically by including noise from sub-threshold modes, as discussed in Sec. VI.B and in Sec. VIII.) Sec. VI–VII also present several other model calculations that illustrate the differences between N-SALT and previous linewidth theories. Finally, in Sec. VIII, we discuss some potential additional corrections that will be addressed in future work. In a second manuscript [41], we also compare the theory against full time-dependent integration of the stochastic Maxwell–Bloch equations and find excellent quantitative agreement with the major results presented here.

Laser dynamics are surveyed in many sources [1–5], but it is useful to review here a simple physical picture of linewidth physics. A resonant cavity [e.g., light bouncing between two mirrors or a photonic-crystal (PhC) microcavity as in Fig. 1(a)] traps light for a long time in some volume, and lasing occurs when a gain medium is “pumped” to a population “inversion” of excited states to the point (*threshold*) where gain balances loss. [Of course, this simple picture is modified once additional modes reach threshold, or for lasers (such as random lasers [42, 43]) in which the passive cavity possesses no strong resonances; all of these complexities are handled by SALT [19, 20] and hence are incorporated into our approach.] For simplicity, consider here a laser operating in the single-mode regime. Above threshold, the gain depends nonlinearly on the mode intensity $|a|^2$, as sketched

in Fig. 1(b): increasing the field intensity decreases the gain due to depletion of the excited states until it reaches a stable steady-state value a_0^2 . (This gain-saturation effect is called “spatial hole-burning” [4] since it can be spatially inhomogeneous.) In the absence of noise, this results in a stable sinusoidal oscillation with an infinitesimal linewidth, but the presence of noise, which can be modeled by random current fluctuations \mathbf{J} [10, 29, 44], perturbs the mode as depicted in Fig. 1(c), resulting in a finite linewidth. There are various sources of noise in real lasers, but spontaneous emission sets a fundamental lower limit on the linewidth [4]; here we will include only spontaneous emission and thermal noise. In particular, although the squared amplitude is stabilized around a_0^2 by the nonlinear gain, the phase ϕ of the mode drifts according to a random walk (a Brownian/Wiener phase) with variance $\langle \phi^2 \rangle \approx \Gamma t$, and the Fourier transform of a Wiener phase yields a Lorentzian lineshape [Fig. 1(d)] with full width at half maximum (FWHM) Γ [28]. The goal of linewidth theory is to derive Γ , ideally given only the thermodynamic FDT description of the current fluctuations and the Maxwell–Bloch physics of the laser cavity.

The most basic approximation for the linewidth (sufficiently far above threshold), usually referred to as the Schawlow-Townes (ST) formula [6, 7], takes the form

$$\Gamma = \frac{\hbar\omega_0\gamma_0^2}{2P}, \quad (1)$$

where P is the output power of the laser, γ_0 is the passive cavity resonance width, and ω_0 is the laser frequency, often approximated to be equal to the real part of the passive-cavity resonance pole at $\omega_* = \omega_0 - i\gamma_0/2$. (A slightly more accurate approximation for the laser frequency takes into account the small line-pulling of the laser frequency towards the atomic transition frequency [45].) The inverse-power dependence causes the famous line-narrowing of a laser above threshold.

Over the decades, a number of now-standard corrections to this formula were found [3–5], leading to the modified ST formula:

$$\Gamma = \frac{\hbar\omega_0\gamma_0^2}{2P} \cdot n_{\text{sp}} \cdot \left| \frac{\int_C dx |\mathbf{E}_c|^2}{\int_C dx \mathbf{E}_c^2} \right|^2 \cdot \left(\frac{\gamma_{\perp}}{\gamma_{\perp} + \frac{\gamma_0}{2}} \right)^2 \cdot (1 + \alpha_0^2). \quad (2)$$

First, the gain medium can be thought of, in many respects, as a system at negative temperature T [46], with the limit of complete inversion of the two lasing levels corresponding to $T \rightarrow 0^-$. When only partial inversion is present, the linewidth is enhanced by a factor of $n_{\text{sp}} \equiv \frac{N_2}{N_2 - N_1}$ [47, 48], where N_2 and N_1 are the *spatially averaged* populations in the upper and lower states of the lasing transition. We refer to this correction as the incomplete-inversion factor (also known as “the spontaneous emission factor”). Second, due to the openness of the laser system, the modes are not power-orthogonal

and the noise power which goes into each lasing mode is enhanced [49]; this correction is known as the Petermann factor, and it becomes significant in low- Q laser systems, where it is not a good approximation to treat the lasing mode \mathbf{E}_c as purely real. ($Q \equiv \omega_0/\gamma_0$ is a dimensionless passive-cavity lifetime defined in units of the optical period [33].) Note that \mathbf{E}_c is the passive-cavity mode [in contrast to SALT solutions, which are the modes of the full non-linear equations, introduced in (6)]. $\int_{\mathcal{C}} dx$ denotes integration over the cavity region. Third, for low- Q laser cavities, it is possible that the gain linewidth γ_{\perp} can be on the order of or smaller than the passive cavity resonance width γ_0 , causing significant dispersion effects as the gain is increased to threshold [9]. This correction is commonly called the “bad-cavity” factor [10, 50]. Unlike the other corrections mentioned above, the bad-cavity factor decreases the laser linewidth. However, very few lasers systems are in the parameter regime where this effect is significant [51]. Finally, amplitude fluctuations in the laser field couple to the phase dynamics, leading to a correction known as the “ α factor”. For atomic gain media, this effect was identified by Lax [9] in the 1960’s, and for this case it is typically a small correction. For bulk semiconductor gain media the effect is large, and typically dominates the broadening due to direct phase fluctuations [52–54]; in this context it is known as the “Henry α factor” [11].

Previous linewidth derivations have taken a number of different approaches, making severe approximations compared to the solution of the full three-dimensional space-dependent Maxwell–Bloch equations in the presence of noise. Generally speaking, linewidth theories can be classified into two categories. The first class includes methods which solve Maxwell’s equations with a phenomenological model for the gain medium and account for noise spatial and spectral correlations by using the FDT [10, 29, 44]. Typically, these methods do not handle nonlinear spatial hole-burning above threshold or multimode effects. These methods, commonly used in the semiconductor laser literature, resulted in linewidth formulas which included the Petermann [49], bad-cavity [2, 10], incomplete-inversion [29], and α factors [11]. Most notably, an early work by Arnaud [55] derived a single-mode linewidth formula without making any simplifying assumptions about the field patterns, handling anisotropic, inhomogeneous, and dispersive media. However, this theory was only applied to very simple, effectively one-dimensional, homogeneous systems, and it was missing hole-burning effects and the α factor.

The second class of linewidth theories consists of scattering-matrix methods [13, 14, 56, 57], which can treat arbitrary geometries without phenomenological parameters and take into account the effects of spatial hole-burning. S-matrix theories only have access to the input and output fields and, therefore, can only treat the noise in a spatially averaged manner and are not able to obtain the α factor rigorously. However, they obtain all of the other corrections to the single-mode linewidth. In

particular, the recent S-matrix approach by Chong *et al.* [13, 14] takes advantage, as we do, of the *ab-initio* computational approach of SALT, and hence has the potential to treat arbitrary geometries and spatial hole-burning effects. (We reduce our results to the most recent scattering-matrix linewidth formula [14] in appendix D.) Note that in practice, S-matrix methods require a substantial independent calculation beyond SALT to extract the linewidths, whereas our approach obtains the linewidths immediately from SALT calculations (or any other method to obtain the steady-state lasing modes) by simple integrals over the fields.

Our derivation of N-SALT, being based on the SALT solutions, has a similar regime of validity. For *single-mode lasing*, SALT and N-SALT are essentially exact, relying only on the rotating-wave approximation and on the laser being sufficiently far above threshold. For multimode lasing, those theories require two additional dynamical constraints [19, 20]: the rates associated with population dynamics must be small compared to both the dephasing rate of the polarization and the lasing mode spacing (roughly, the free spectral range). The former constraint is satisfied in all solid-state lasers, whereas the latter requires a sufficiently small laser cavity. The actual size depends both on details of the cavity and of the gain medium used, but the appropriate limit is realized in many complex lasers of interest. When these frequency scales are not well-separated, the level populations are not quasi-stationary, and multimode SALT will initially lose accuracy and eventually fail completely (since multimode lasing becomes unstable [58]). Moreover, while the average (SALT) behavior is unaffected by non-lasing poles, they do affect the noise properties, and N-SALT in its current form only accounts for a finite number of poles in the Green’s function (appendix A.2). [We only include lasing poles (i.e., poles on the real axis), but extension to include non-lasing poles, which determine the amplified spontaneous emission (ASE) [40, 59], will be straightforward (Sec. VIII)]. As noted above, the linewidth formula additionally assumes that the laser is operating far enough above threshold that amplitude fluctuations are small compared to the steady state amplitudes (i.e., $|a(t)| \approx a_0$ in the notation of Sec. V). Hence, our formula does not describe the linewidth near the lasing thresholds. Our perturbation approach takes into account only the lowest-order correction to the complex modal amplitude $a(t)$ and neglects higher-order corrections to the frequency ω_0 and spatial pattern $\mathbf{E}_0(\mathbf{x})$ [see Eq. (7)]. Moreover, we neglect non-Lorentzian corrections to the lineshape [60–64] (Sec. IV). In the following section we present our generalized linewidth formula in the single-mode regime (3) and compare it with traditional linewidth theories.

II. THE N-SALT LINEWIDTH FORMULA

Our main result is a multimode linewidth formula which generalizes (2). In the multimode case, the result takes the form of a covariance matrix for the phases of the various modes, which is presented in (36,37) of Sec. V. In the single-mode case, the N-SALT linewidth formula takes the simple form:

$$\Gamma = \frac{\hbar\omega_0\tilde{\gamma}_0^2}{2P} \cdot \tilde{n}_{\text{sp}} \cdot \tilde{K} \cdot \tilde{B} \cdot (1 + \tilde{\alpha}^2). \quad (3)$$

The modified correction factors (marked by tildes) are defined in Table. I. As can be seen from the table, those factors generalize the traditional expressions by taking into account both spatial inhomogeneity and nonlinearity. Since the generalized factors depend on the SALT permittivity ε , mode profile $\mathbf{E}_0(\mathbf{x})$, and frequency ω_0 , one can no longer regard the effects of cavity-openness, nonlinearity, and dispersion as separate multiplicative effects. In this sense, our formula demonstrates the intermingled nature of the linewidth correction factors, as previously introduced in [13, 14], but here demonstrated in a new level of generality. We denote by $\int dx$ integration over all space, for any number of spatial dimensions. We use the shorthand notation for vector products $|\mathbf{E}_0|^2 = \mathbf{E}_0 \cdot \mathbf{E}_0^*$ and $\mathbf{E}_0^2 = \mathbf{E}_0 \cdot \mathbf{E}_0$, where the latter unconjugated inner product appears naturally because of the biorthogonality relation for lossy complex-symmetric systems [65, 66]. $\text{Im } \varepsilon(\mathbf{x})$ denotes the imaginary part of the nonlinear steady-state permittivity (5), which is negative/positive in gain/loss regions. The output power P is related to the SALT solutions by invoking Poynting's theorem, which one can use to show that $P \propto \int_P dx [-\text{Im } \varepsilon(\mathbf{x})] |\mathbf{E}_0(\mathbf{x})|^2$. We use $\int_P dx$ to denote some volume which contains the gain medium. The choice of the volume is somewhat arbitrary; e.g., integrating over the cavity region corresponds to the output power at the cavity boundary [29]. Note, however, that this arbitrariness in the choice of the volume is not a general feature of our formula. After substituting the relevant expressions from Table. I into (3), the integrals which contain $\int_P dx$ cancel, resulting in an expression for the linewidth only in terms of integrals over the entire space. The effective inverse temperature $\beta(\mathbf{x})$ is determined by the inhomogeneous steady-state atomic populations $N_1(\mathbf{x})$ and $N_2(\mathbf{x})$, and is defined as [67–69]

$$\beta(\mathbf{x}) \equiv \frac{1}{\hbar\omega_0} \ln \left(\frac{N_1(\mathbf{x})}{N_2(\mathbf{x})} \right). \quad (4)$$

In regions where the gain medium is pumped sufficiently to invert the population, $\beta(\mathbf{x})$ is negative; in regions where the pump is too weak to invert, $\beta(\mathbf{x})$ will be positive [and still given by (4)]; and in unpumped regions, Eq. (4) will simply reduce to the equilibrium temperature of the surrounding environment $(k_B T)^{-1}$. The quantities $N_1(\mathbf{x})$ and $N_2(\mathbf{x})$ are an output of the SALT solution in the absence of noise. The spatially dependent expression inside the square brackets in the definition of \tilde{n}_{sp}

Symbol	Traditional	Generalized
$\tilde{\gamma}_0$ cavity decay rate	γ_0	$\left \frac{\int dx (\omega_0 \text{Im } \varepsilon) \mathbf{E}_0^2}{\int dx \varepsilon \mathbf{E}_0^2} \right $
\tilde{n}_{sp} incomplete inversion	$\frac{N_2}{N_2 - N_1}$	$\frac{\int dx \left[\frac{1}{2} \coth\left(\frac{\hbar\omega\beta}{2}\right) - \frac{1}{2} \right] \text{Im } \varepsilon \mathbf{E}_0 ^2}{\int_P dx \text{Im } \varepsilon \mathbf{E}_0 ^2}$
\tilde{K} Petermann	$\left \frac{\int_C dx \mathbf{E}_c ^2}{\int_C dx \mathbf{E}_c^2} \right ^2$	$\left \frac{\int_P dx \text{Im } \varepsilon \mathbf{E}_0 ^2}{\int dx \text{Im } \varepsilon \mathbf{E}_0^2} \right ^2$
\tilde{B} bad cavity	$\left(\frac{\gamma_{\perp} + \frac{\gamma_0}{2}}{\gamma_{\perp}} \right)^2$	$\left \frac{\int dx \varepsilon \mathbf{E}_0^2}{\int dx \mathbf{E}_0^2 \left(\varepsilon + \frac{\omega_0}{2} \frac{\partial \varepsilon}{\partial \omega_0} \right)} \right ^2$
$\tilde{\alpha}$ amplitude-phase coupling	$\frac{\omega_a - \omega_0}{\gamma_{\perp}}$	$\frac{\text{Im } C}{\text{Re } C}$
C nonlinear coupling		$\frac{-i\omega_0 \int dx \frac{\partial \varepsilon}{\partial \mathbf{a} ^2} \mathbf{E}_0^2}{\int dx \left(\varepsilon + \frac{\omega_0}{2} \frac{\partial \varepsilon}{\partial \omega_0} \right) \mathbf{E}_0^2}$

TABLE I. Traditional and new linewidth correction factors for the single-mode linewidth formulas (2,3).

in Table. I generalizes the spatially averaged incomplete-inversion factor $\frac{N_2}{N_2 - N_1}$. That can be seen by noting that $\frac{1}{2} \coth\left(\frac{\hbar\omega\beta}{2}\right) - \frac{1}{2} = (\exp[\hbar\omega\beta] - 1)^{-1} \equiv n_B$, where n_B is the usual Bose–Einstein distribution function [70, 71]. (For gain media, it is sometimes convenient to introduce the positive spontaneous-emission factor $n_{\text{sp}} = -n_B$ [72]. Note that this definition ensures that the generalized incomplete-inversion factor is always positive.) The $\frac{1}{2}$ factor subtracted from the hyperbolic cotangent was discussed in [72], and we give a simple classical explanation for it in appendix E. If standard absorbing layers are used to implement outgoing boundary conditions in the SALT solver [21] and the temperature of the ambient medium is assigned to these layers, then the N-SALT formula includes the effect of incoming thermal radiation. A generalized Petermann factor which formally resembles \tilde{K} appeared in previous work by Schomerus [57] (in his expression for the Petermann factor for TM modes in two-dimensional dielectric resonators). However, the earlier formula is expressed in terms of passive resonance scalar fields, whereas our correction contains 3d nonlinear SALT solutions. Finally, $\tilde{\alpha}$ is a generalized α factor, defined explicitly in Sec. V (30). For atomic gain media, the traditional factor is expressed in terms of the atomic transition frequency ω_a and decay rate of the atomic polarization γ_{\perp} . In the current work we will only evaluate the atomic case, although the general expression in terms of the non-linear coupling C should also apply to the semiconductor case.

The N-SALT formula (3) reduces to the traditional formula (2) in some limiting cases. Let us consider, for simplicity, a 1d Fabry-Pérot laser cavity of length L surrounded by air (i.e., $\text{Im } \varepsilon = 0$ outside the cavity region). Let us assume also that the laser is operating not too far above the threshold and is uniformly pumped, hence $\text{Im } \varepsilon$ and β are nearly constant inside the cavity. In this limit, all the integrals in Table. I can be approximated by reducing the integration limits to the cavity region; terms which contain integration over the imaginary part of the permittivity are non-zero only within the cavity region (e.g., $\int dx \text{Im } \varepsilon |\mathbf{E}_0|^2$ becomes $\text{Im } \varepsilon \int_C dx |\mathbf{E}_0|^2$); while terms of the form $\int dx \varepsilon \mathbf{E}_0^2$ can be written as the sum of the cavity contribution $\varepsilon \int_C dx \mathbf{E}_0^2$ and the surrounding medium contribution $\int_{\text{out}} dx \mathbf{E}_0^2$, where the latter is negligible for $L\omega_0 \gg 1$, as shown in appendix D and in [14] (here and throughout the paper, we are setting $c = 1$). Using this approximation, it is immediately apparent from Table. I that the incomplete-inversion factor reduces to the traditional expression. The generalized Petermann factor reduces to the traditional factor in the limit of a high-Q cavity, where the threshold lasing state $\mathbf{E}(\mathbf{x})$ is approximately the same as the passive resonance state $\mathbf{E}_c(\mathbf{x})$. In order to simplify the remaining terms, recall that the lasing threshold is reached when gain in the system compensates for the loss. For weak losses (small $\text{Im } \varepsilon / \varepsilon$) that can be treated by perturbation theory, the threshold condition is $\gamma_0 = \frac{\omega_0 \text{Im } \varepsilon}{\varepsilon}$ [2] and, therefore, the generalized decay rate reduces to γ_0 (one can thereby see that the Schawlow-Townes formula (2) neglects nonlinear corrections to γ_0 , as was also shown in [13]). Next, let us discuss the generalized bad-cavity factor, which simplifies to $\left(1 + \frac{\omega_0}{2\varepsilon} \frac{\partial \varepsilon}{\partial \omega_0}\right)^{-2}$ after reducing the integration limits. In order to show that it agrees with the traditional factor, we need to show that $\frac{\omega_0}{2\varepsilon} \frac{\partial \varepsilon}{\partial \omega_0} \approx \frac{\gamma_0}{2\gamma_\perp}$. The steady-state effective permittivity, as used in SALT theory (appendix A.1), is

$$\varepsilon(\mathbf{x}) = \varepsilon_c(\mathbf{x}) + \frac{\gamma_\perp D(\mathbf{x})}{\omega_0 - \omega_a + i\gamma_\perp}, \quad (5)$$

where ε_c is the passive permittivity and the second term is the active nonlinear permittivity due to the gain medium. The population inversion $D(\mathbf{x}) = N_2(\mathbf{x}) - N_1(\mathbf{x})$ is generally spatially varying above threshold due to spatial hole-burning. Since we assume here that we are close to threshold and that the pumping is uniform, the inversion is also uniform in space and near its threshold value. If one assumes, additionally, that the detuning of the lasing frequency from atomic resonance is small ($|\omega_0 - \omega_a| \ll \gamma_\perp$), one obtains $\frac{\partial \varepsilon}{\partial \omega_0} \approx \frac{\text{Im } \varepsilon}{\gamma_\perp}$. Finally, we show in Sec. VI.A that our $\tilde{\alpha}$ reduces to the known α_0 in homogeneous low-loss cavities, so that all factors of the corrected ST formula are recovered in this limit. (Note that line-pulling effects which may modify the lasing frequency ω_0 are handled by SALT.)

In the next section, we present the TCMT equations which are used in this paper to derive the N-SALT

linewidth formula (3), but which may also be used to extract more information on laser dynamics away from steady state.

III. THE N-SALT TCMT EQUATIONS

In the absence of noise, the electric field of a laser operating in the multimode regime is given by the real part of $\mathbf{E}_0(\mathbf{x}, t)$, where

$$\mathbf{E}_0(\mathbf{x}, t) = \sum_{\mu} \mathbf{E}_{\mu}(\mathbf{x}) a_{\mu 0} e^{-i\omega_{\mu} t}, \quad (6)$$

and the laser has zero linewidth. (This assumes, of course, that there exists a steady-state multimode solution of the nonlinear semi-classical lasing equations [19, 20].) The modes $\mathbf{E}_{\mu}(\mathbf{x})$ and frequencies ω_{μ} can be calculated using SALT, which solves the semi-classical Maxwell-Bloch equations in the absence of noise. (SALT has been generalized to include multi-level atoms [73], multiple lasing transitions, and gain diffusion [74]; any of these cases can thus be treated by N-SALT with minor modifications, but we focus on the two-level case here.) The linewidth can now be calculated by adding Langevin noise, as described below.

In the presence of a weak noise source, the electric field can be written as a superposition of the steady-state lasing modes with time-dependent amplitudes $a_{\mu}(t)$ which fluctuate around $a_{\mu 0}$:

$$\mathbf{E}(\mathbf{x}, t) = \sum_{\mu} \mathbf{E}_{\mu}(\mathbf{x}) a_{\mu}(t) e^{-i\omega_{\mu} t}. \quad (7)$$

In principle, the sum in (7) should also include the non-lasing modes since the set of lasing modes by itself does not form a complete basis for the fields. Non-lasing modes contribute to amplified spontaneous emission (ASE), which has a significant effect on the spectrum near and below the lasing thresholds [40, 59] and will be treated in future work.

In appendix A, we derive the N-SALT TCMT equations of motion for $a_{\mu}(t)$ starting with the full vectorial Maxwell-Bloch equations. We show that the noise-driven field obeys an effective nonlinear equation which, in the frequency domain, takes the form

$$[\nabla \times \nabla \times - \omega^2 \varepsilon(\omega, a)] \hat{\mathbf{E}}(\mathbf{x}, \omega) = \hat{\mathbf{F}}_S(\mathbf{x}, \omega), \quad (8)$$

where the carets denote Fourier transforms [e.g., $\mathbf{E}(\mathbf{x}, t) \equiv \int_0^\infty d\omega e^{-i\omega t} \hat{\mathbf{E}}(\mathbf{x}, \omega)$]. Spontaneous emission is included via the stochastic noise term $\hat{\mathbf{F}}_S(\mathbf{x}, \omega)$ (quantified in Sec. IV), and the effective permittivity $\varepsilon(\omega, a)$ (derived in appendix A.2) is given by

$$\varepsilon(\omega, a) \hat{\mathbf{E}}(\mathbf{x}, \omega) = \sum_{\mu} \left[\varepsilon_c \hat{a}_{\mu} + \frac{\gamma_\perp}{\omega - \omega_a + i\gamma_\perp} \hat{D} * \hat{a}_{\mu} \right] \mathbf{E}_{\mu}(\mathbf{x}), \quad (9)$$

where the asterisk denotes a convolution. The second argument of $\varepsilon(\omega, a)$ denotes the implicit dependence of ε on the modal amplitudes a_μ through the inversion \hat{D} . The effective permittivity (9) can be decomposed into a steady-state-amplitude dispersive term and a nonlinear non-dispersive term (similar in spirit to [75]). The key point here is that, to lowest order, there are two corrections to the permittivity in the presence of noise: the *dispersive correction* due to any shift in frequency at the unperturbed amplitudes $a_{\mu 0}$, and the *nonlinear correction* due to any shift in amplitude at the unperturbed frequency. (Shifts in frequency are small because only frequency components within the mode linewidths matter, while shifts in amplitude are small because of the stabilizing effect of gain feedback.) The coupling between these two perturbations is higher order and is hence dropped, which greatly simplifies the analysis.

Substituting the permittivity expansion (derived explicitly in appendix A.3) into Maxwell's equation (8), we find that the noise-driven field obeys the linearized equation

$$[\nabla \times \nabla \times -\omega^2 \varepsilon(\omega, a_0)] \hat{\mathbf{E}}(\mathbf{x}, \omega) = \hat{\mathbf{F}}_{\text{NL}}(\mathbf{x}, \omega) + \hat{\mathbf{F}}_{\text{S}}(\mathbf{x}, \omega), \quad (10)$$

i.e., the dispersive permittivity which appears on the left-hand side of (10) is evaluated at the steady-state amplitude a_0 . The nonlinear non-dispersive term $\hat{\mathbf{F}}_{\text{NL}}$ [defined explicitly in (A23)], which corresponds to amplitude fluctuations at the unperturbed frequency, appears as a restoring force on the right-hand side. The noise-driven field $\hat{\mathbf{E}}(\mathbf{x}, \omega)$ is found in appendix A.4 by convolving the linearized Green's function with the source terms $\hat{\mathbf{F}}_{\text{NL}}$ and $\hat{\mathbf{F}}_{\text{S}}$. Finally, the N-SALT TCMT equations are obtained by transforming the noise-driven field back into the time domain.

A. Time-delayed multimode model

We find that, in the most general case, the TCMT equations take the form

$$\dot{a}_\mu = \sum_\nu \int dx c_{\mu\nu}(\mathbf{x}) \times \left[\gamma(\mathbf{x}) \int_0^t dt' e^{-\gamma(\mathbf{x})(t-t')} (a_{\nu 0}^2 - |a_\nu(t')|^2) \right] a_\mu + f_\mu. \quad (11)$$

Comparing (11) and (10), one can see that the first term on the right-hand side of (11) is related to the nonlinear restoring force $\hat{\mathbf{F}}_{\text{NL}}$, and the Langevin noise $f_\mu(t)$ is associated with $\hat{\mathbf{F}}_{\text{S}}$.

The nonlinear coupling coefficients $c_{\mu\nu}(\mathbf{x})$ [derived in (A33)] correspond to local changes in the nonlinear permittivity with respect to intensity changes in each of the

modes

$$c_{\mu\nu} = \frac{-i\omega_\mu^2 \frac{\partial \varepsilon(\omega_\mu)}{\partial |a_\nu|^2} \mathbf{E}_\mu^2}{\int dx (\omega_\mu^2 \varepsilon)'_\mu \mathbf{E}_\mu^2}, \quad (12)$$

where we have introduced a shorthand notation for the derivative in the denominator $(\omega_\mu^2 \varepsilon)'_\mu \equiv \frac{\partial}{\partial \omega} \omega^2 \varepsilon|_{\omega_\mu}$. This modal coupling in the fluctuation dynamics comes about because of saturation of the gain: a fluctuation in mode μ affects the amplitudes of all the other modes ν .

The N-SALT TCMT equations are nonlocal in time because the atomic populations are not in general able to follow the field fluctuations instantaneously and, instead, respond with a time delay determined by the local atomic decay rate $\gamma(\mathbf{x})$, given by

$$\gamma(\mathbf{x}) = \gamma_{\parallel} \left(1 + \sum_\nu \frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |a_{\nu 0}|^2 |\mathbf{E}_\nu|^2 \right). \quad (13)$$

The second term in (13) is precisely the local enhancement of the atomic decay rate due to stimulated emission in the presence of the lasing fields. (A simplified *spatially averaged* enhancement of the atomic decay rate was previously discussed in [76].)

The Langevin force f_μ is the projection of the spontaneously emitted field onto the corresponding mode \mathbf{E}_μ [29]. Defining $\mathbf{F}_\mu(t) \equiv \mathbf{F}_\text{S} e^{i\omega_\mu t}$, the Langevin force f_μ is

$$f_\mu(t) = \frac{i \int dx \mathbf{E}_\mu \cdot \mathbf{F}_\mu(t)}{\int dx (\omega_\mu^2 \varepsilon)'_\mu \mathbf{E}_\mu^2}. \quad (14)$$

The full N-SALT TCMT equations (11) describe the most typical situation in laser dynamics of a “class B” laser [35–37], in which the polarization of the gain medium can be adiabatically eliminated but the population dynamics is relatively slow and cannot be so eliminated. However, much of the basic linewidth physics can be extracted from the limit when the population dynamics is also adiabatically eliminable, which describes “class A” lasers. Since the mathematical analysis is simpler in this limit, we will begin the spectral analysis in Sec. V with the latter model. We discuss this limit, which we refer to as the “instantaneous model,” in the following section.

B. Instantaneous single-mode model

When the population relaxation rate $\gamma(\mathbf{x})$ is (everywhere) large compared to the dynamical scales determining $a_\mu(t)$, the exponential terms in (11) act like δ functions. After the spatial integration, and specializing in this section to the single-mode case, we obtain the simple nonlinear oscillator model driven by a weak Langevin

force $f(t)$:

$$\dot{a} = C (a_0^2 - |a|^2) a + f, \quad (15)$$

where $C = \int dx c(\mathbf{x})$ is the integrated nonlinear coupling. This instantaneous nonlinear oscillator model was previously introduced by Lax [9, 28], and has been used extensively in linewidth theories [2]. The N-SALT approach enables computing the model's parameters *ab initio*, taking full account of the spatial hole-burning term and the vectorial nature of the fields [including multimode effects, when generalizing (15) to the multimode regime]. Also, our approach shows that this well-known model can be explicitly derived from the more general (non-instantaneous) model, presented in the previous section. Above the lasing threshold, $a_0 > 0$ and $\text{Re}[C] > 0$, and the system undergoes self-sustained oscillations with a stable steady state at $|a| = a_0$, as demonstrated in Fig. 1(b). In fact, near threshold one can show that $\text{Re}[C]$ is approximately the threshold gain, which balances the cavity loss κ . Hence the dynamical scale of $a(t)$ is of order κ , which must then be much smaller than $\gamma(\mathbf{x})$ for the instantaneous model to hold; this is the standard dynamical condition for class A lasers [35–37].

The nonlinear term in (15) and the multimode counterpart in (12) are derived rigorously in appendix A, but we can motivate the resulting expressions using simple physical arguments. The nonlinear term can be viewed as a shift in the oscillation frequency, i.e., $-i\Delta\omega = C(a_0^2 - |a|^2)$. Using first-order perturbation theory [77], the frequency shift due to a change in dielectric permittivity $\Delta\epsilon$ is given by

$$\Delta\omega = -\omega_0^2 \frac{\int dx \Delta\epsilon \mathbf{E}_0^2}{\int dx (\omega_0^2 \epsilon)'_0 \mathbf{E}_0^2}. \quad (16)$$

Plugging in the differential of the permittivity due to small changes in the squared mode amplitude, $\Delta\epsilon \approx \frac{\partial\epsilon}{\partial|a|^2}(|a|^2 - a_0^2)$, we find that the coupling coefficient in the instantaneous model is

$$C = \frac{-i\omega_0^2 \int dx \frac{\partial\epsilon}{\partial|a|^2} \mathbf{E}_0^2}{\int dx (\omega_0^2 \epsilon)'_0 \mathbf{E}_0^2}. \quad (17)$$

This is the single-mode version of (12) integrated over space due to rapid relaxation. As we will see, this simple result, combined with the spectrum of the Langevin noise (section IV), is all that is needed to derive the single-mode linewidth formula (3) (see Section V), and the multimode generalization also follows straightforwardly. Hence, after analyzing the noise spectrum, we will first derive the linewidth within the instantaneous model before moving on to the more complicated case of the full N-SALT TCMT equations. The latter will show that the basic linewidth formula is unchanged from that of the instantaneous model except for the addition of side peaks due to the relaxation oscillations present in class B lasers.

IV. THE AUTOCORRELATION FUNCTION OF THE LANGEVIN FORCE

In this section, we express the autocorrelation function of the Langevin force f_μ

$$\langle f_\mu(t) f_\nu^*(t') \rangle = R_\mu \delta_{\mu\nu} \delta(t - t') \quad (18)$$

in terms of the autocorrelation function of the noise source \mathbf{F}_μ . It is well known that quantum and thermal fluctuations can be modeled as zero-mean random variables, defined by their correlation functions [26, 27]. This Rytov picture [24] is essentially a consequence of the central-limit theorem (CLT) [78, 79], which holds since the classical forcing \mathbf{F}_S is the sum of a large number of randomly emitted photons. The autocorrelation function of \mathbf{F}_S can be found by invoking the fluctuation-dissipation theorem (FDT), as explained below.

The probability distributions of the pumped medium and the electromagnetic field obey Boltzmann statistics, with an effective local temperature β defined in terms of the atomic inversion [46] (see definition in Sec. II). Under the typical conditions of local thermal equilibrium [23–27], dissipation by optical absorption must be balanced by spontaneous emission from current fluctuations $\mathbf{J}(\mathbf{x}, t)$. One can then apply the FDT for the Fourier-transformed forcing $\hat{\mathbf{F}}_S(\mathbf{x}, \omega) = -4\pi i\omega \hat{\mathbf{J}}(\mathbf{x}, \omega)$ [80]:

$$\begin{aligned} \langle \hat{\mathbf{F}}_S(\mathbf{x}, \omega) \hat{\mathbf{F}}_S^*(\mathbf{x}', \omega') \rangle = \\ 2\hbar\omega^4 \text{Im} \epsilon(\mathbf{x}, \omega) \coth\left(\frac{\hbar\omega\beta(\mathbf{x})}{2}\right) \delta(\mathbf{x} - \mathbf{x}') \delta(\omega - \omega'). \end{aligned} \quad (19)$$

Using this result, we calculate the autocorrelation of the Langevin force \hat{f}_μ [i.e., the Fourier transform of (14), defined as $\hat{f}_\mu(\omega) \equiv \frac{1}{2\pi} \int_0^\infty dt e^{i\omega t} f_\mu(t)$] and we obtain

$$\langle \hat{f}_\mu(\omega) \hat{f}_\nu^*(\omega') \rangle = \hat{R}_\mu(\omega) \delta(\omega - \omega') \delta_{\mu\nu}, \quad (20)$$

where the frequency-domain autocorrelation coefficient is

$$\hat{R}_\mu(\omega) = 4\hbar\omega^4 \frac{\int dx |\mathbf{E}_\mu|^2 \text{Im} \epsilon(\omega) \left[\frac{1}{2} \coth\left(\frac{\hbar\omega\beta}{2}\right) - \frac{1}{2} \right]}{\left| \int dx \mathbf{E}_\mu^2 (\omega_\mu^2 \epsilon)'_\mu \right|^2}. \quad (21)$$

The $\frac{1}{2}$ factor subtracted from the hyperbolic cotangent is explained in appendix E and in [72].

The time-domain diffusion coefficient R_μ can be found directly from (21) taking the inverse Fourier transform. For the common case of a small linewidth, $\text{Im} \epsilon(\omega)$ and $\coth\left(\frac{\hbar\omega\beta}{2}\right)$ are nearly constant for frequencies within the linewidth. [This means, essentially, that the Langevin force $f_\mu(t)$ can be treated as white noise]. Consequently, one can approximate the diffusion coefficient in (21) by its value at ω_μ . With this simplification, the time-domain

diffusion coefficient in (18) is conveniently given by $R_\mu = 2\pi\hat{R}_\mu(\omega_\mu)$ [29].

More generally, however, including this frequency dependence corresponds to temporally correlated fluctuations, leading to non-Lorentzian corrections to the laser lineshape [60–64]. These “memory effects” can be addressed using our approach (as discussed in Sec. VIII) and we plan to include them in future work.

V. THE LASER SPECTRUM

In this section, we calculate the laser spectrum using the N-SALT TCMT equations (11,15) and the noise autocorrelation function (20,21). We begin by showing that the phase of the lasing mode undergoes simple Brownian motion; consequently, the laser spectrum is a Lorentzian, with a width given by the phase-diffusion coefficient. In Sec. V.A, we calculate the phase-diffusion coefficient (hence the linewidth) for the instantaneous model (15) and in Sec. V.B, we outline the analysis for the time-delayed model (11), leaving the details of the derivation to appendix B. More accurately, the spectrum of the time-delayed model consists of a central Lorentzian peak at the lasing resonance frequency and additional side peaks due to relaxation oscillations, which are present in class B lasers. The latter side peaks are the subject of Sec. V.C.

A. Instantaneous single-mode model

The complex mode amplitude $a(t)$ can be written in polar form as

$$a(t) = [a_0 + \delta(t)] e^{i\phi(t)}. \quad (22)$$

a_0 is the steady-state amplitude, while δ and ϕ are *real* amplitude and phase fluctuations. Substituting the modal expansion (22) in (15), defining

$$\begin{aligned} A &\equiv 2a_0^2 \text{Re } C \\ B &\equiv 2a_0^2 \text{Im } C, \end{aligned} \quad (23)$$

and keeping terms to first order in δ/a_0 , we obtain

$$\dot{\delta} = -A\delta + f_R, \quad (24)$$

$$a_0\dot{\phi} = -B\delta + f_I, \quad (25)$$

where $f_R \equiv \text{Re}\{f\}$ and $f_I \equiv \text{Im}\{f\}$. We check the approximation of $|\delta| \ll a_0$ *a posteriori* and we find that it generally holds (as was also shown in [40]), except near threshold ($a_0 \rightarrow 0$), which is a case we discuss in Sec. VI.C.

When the nonlinear coupling coefficient is real ($B = 0$), it is evident from (25) that the phase undergoes simple Brownian motion (i.e., it is a Wiener process) and hence

the phase variance increases linearly in time. An oscillator with Brownian phase noise has a Lorentzian spectrum [81], and one can reproduce that result briefly as follows. The laser spectrum $S_a(\omega)$ is given by the Fourier transform of the autocorrelation function of $a(t)$:

$$\langle a(t)a^*(0) \rangle \approx a_0^2 \langle e^{-i(\phi(t)-\phi(0))} \rangle = a_0^2 e^{-\frac{1}{2}\langle (\phi(t)-\phi(0))^2 \rangle}. \quad (26)$$

For a Wiener phase, whose variance is $\langle (\phi(t) - \phi(0))^2 \rangle = \Gamma|t|$, the Fourier transform of (26) is a Lorentzian whose central-peak width is Γ [28]. In passing from the first to second step in (26), one neglects direct amplitude-fluctuation contributions (which are decoupled from the phase) as these only introduce broad-spectrum background noise, but do not affect the linewidth of the laser peak (we return to this point in Sec. V.C). In passing from the second to the third step, one assumes that the phase is a Gaussian normal variable, which is justified as a consequence of the CLT.

It is well known that also in the general case of $B \neq 0$, the phase is a Wiener process, with a modified diffusion coefficient [11]. In order to calculate the phase variance explicitly, we solve (24,25) and obtain

$$\delta(t) = \int_0^t e^{-A(t-t')} f_R(t') dt', \quad (27)$$

$$a_0\phi(t) = -B \int_0^t \delta(t') dt' + \int_0^t f_I(t') dt'. \quad (28)$$

Substituting (27) into (28), using the autocorrelation function of f (21), and performing the integration, one obtains that the phase variance in the long-time limit is $\langle (\phi(t) - \phi(0))^2 \rangle = \frac{R}{2a_0^2} \left(1 + \left(\frac{B}{A}\right)^2\right) |t|$ (where terms growing more slowly than $|t|$ were neglected, as explained in greater detail in appendix B). Therefore, the linewidth is

$$\Gamma = \frac{R}{2a_0^2} (1 + \tilde{\alpha}^2), \quad (29)$$

where we have defined the generalized α factor:

$$\tilde{\alpha} = \frac{B}{A} = \frac{\text{Im } C}{\text{Re } C}, \quad (30)$$

with the nonlinear coefficient C defined in (17). Substituting the autocorrelation function (21) in (29) and using Poynting’s theorem to relate a_0^2 to the output power $P = \frac{\omega_0 a_0^2}{2\pi} \int_P dx (-\text{Im } \varepsilon(\mathbf{x})) |\mathbf{E}_0(\mathbf{x})|^2$ [82], we obtain the single-mode linewidth formula (3). From (29), it is evident that the Schawlow–Townes, Petermann, bad-cavity and incomplete-inversion factors are all included in the term $\frac{R}{2a_0^2}$, and generally cannot be separated into the traditional factors of (2) [14].

When the nonlinear coupling coefficient is complex (i.e., when $B \neq 0$), the resonance peak is not only broadened but is also shifted [11]. The shift in center frequency

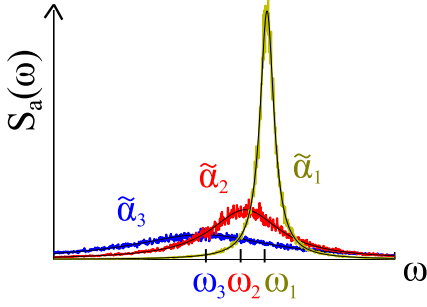


FIG. 2. (Color online) Simulated spectrum $S_a(\omega)$ of the instantaneous model (15) with $\text{Re } C = 10$, noise autocorrelation coefficient $R = 0.1$ and three values of $\tilde{\alpha}$: 10 (blue), 5 (red), 1 (yellow) (C, R, S_a , and ω are given in arbitrary frequency units). The noisy signal is the simulation result and the black curves are Lorentzian lineshapes with widths Γ and center frequency shifts given by (29,31).

is found by keeping second-order terms in δ/a_0 and calculating the average phase drift:

$$\delta\omega = \langle \dot{\phi} \rangle = -\frac{RB}{4a_0^2 A}. \quad (31)$$

An identical formula was derived in [29] in a phenomenological instantaneous model.

Fig. 2 shows the spectrum of the instantaneous model, which is obtained by numerically solving (15) using a stochastic Euler scheme [83]. Introducing the notation $\mathcal{F}(a) \equiv C(a_0^2 - |a|^2)a$ and discretizing time as $a(n\Delta t) \approx a_n$, the Euler update equation for the n -th step is

$$a_n = a_{n-1} + \mathcal{F}(a_{n-1}) \Delta t + \sqrt{R \Delta t} \zeta, \quad (32)$$

where ζ is a gaussian random variable of mean 0 and variance 1, i.e., $\zeta \in N(0,1)$. [For the data presented in Fig. 2, Δt was decreased until the simulation results converged. In later sections (Fig. 3), we implemented a fourth-order Runge–Kutta method in order to achieve convergence]. The simulated spectra (noisy colorful curves) match the predicted Lorentzian lineshapes (solid black curves), which are calculated using (29,31). As $\tilde{\alpha}$ increases, the linewidths are broadened and the center frequencies are shifted.

B. Time-delayed multimode model

We now turn to the laser spectrum produced by the time-delayed model, where the nonlinearity is dependent on the modal amplitudes at previous times. Although we calculate the linewidth of the full time-delayed N-SALT TCMT equations (11) in appendix B, we begin this section by considering the simplified case of a spatially homogeneous medium $\gamma(\mathbf{x}) \approx \gamma_0$ (this is a good

approximation for a uniformly pumped class B laser operating near threshold). In this case, the single-mode time-delayed model takes the form

$$\dot{a} = C \left(\gamma_0 \int^t dt' e^{-\gamma_0(t-t')} (a_0^2 - |a(t')|^2) \right) a + f, \quad (33)$$

where $C = \int d\mathbf{x} c(\mathbf{x})$ is the integrated nonlinear coupling and $c(\mathbf{x})$ is defined in (12). This integro-differential equation can be turned into a first-order ODE by using the modal expansion from Sec. V.A: $a = (a_0 + \delta)e^{i\phi}$, keeping terms to first order in δ/a_0 , and introducing the variable

$$\xi(t) = \gamma_0 \int^t dt' e^{-\gamma_0(t-t')} \delta(t'). \quad (34)$$

Then, (33,34) can be recast in the form $\dot{\mathbf{v}} = \mathbb{K}\mathbf{v} + \mathbf{f}$, where $\mathbf{v} = \{\delta, a_0\phi, \xi\}$.

However, most generally, the spatial dependence of $\gamma(\mathbf{x})$ cannot be neglected. The time-averaged deviation $\xi(\mathbf{x}, t)$ is therefore spatially dependent, and one obtains an infinite-dimensional problem. To simplify the algebra, we discretize space [e.g., discretizing (11) into a Riemann sum over sub-volumes V_k] and recover the continuum limit at the end. This yields the discrete-space multimode model:

$$\begin{aligned} \dot{a}_\mu = & \sum_{\nu k} C_{\mu\nu}^k \left(\gamma_k \int^t dt' e^{-\gamma_k(t-t')} (a_{\nu 0}^2 - |a_\nu(t')|^2) \right) a_\mu + f_\mu, \end{aligned} \quad (35)$$

where the discretized nonlinear coupling coefficients are $C_{\mu\nu}^k = \int_{V_k} d\mathbf{x} c_{\mu\nu}(\mathbf{x})$ (so that $C_{\mu\nu} = \sum_k C_{\mu\nu}^k$), γ_k is the relaxation rate at the k 'th spatial point and $a_{\nu 0}$ is the steady-state amplitude of mode ν .

In appendix B, we study the statistical properties of the solutions to (35). We introduce the M -dimensional vectors whose entries are $\Phi_\mu \equiv a_{\mu 0} \phi_\mu$ (where M is the number of active lasing modes) and we calculate the covariance matrix $\langle \Phi_\mu(t) \Phi_\nu(0) \rangle$. We find that the result is independent of the relaxation rates γ_k or the discretization scheme:

$$\langle \Phi(t) \Phi^T(0) \rangle = \left(\frac{\mathbb{R}}{2} + \mathbb{B} \mathbb{A}^{-1} \frac{\mathbb{R}}{2} (\mathbb{B} \mathbb{A}^{-1})^T \right) |t|. \quad (36)$$

The matrices \mathbb{A} and \mathbb{B} correspond to the real and imaginary parts of the coupling matrices, with entries $A_{\mu\nu} = 2a_{\mu 0} a_{\nu 0} \text{Re}[C_{\mu\nu}]$ and $B_{\mu\nu} = 2a_{\mu 0} a_{\nu 0} \text{Im}[C_{\mu\nu}]$. \mathbb{R} is the autocorrelation function of the Langevin force vector \mathbf{f} [defined in (21)]. The diagonal of this matrix, divided by $|t|$ and by the squared modal amplitude, gives the generalized linewidths

$$\Gamma_\mu = \frac{1}{2a_{\mu 0}^2} \left(R_{\mu\mu} + [\mathbb{B} \mathbb{A}^{-1} \mathbb{R} (\mathbb{B} \mathbb{A}^{-1})^T]_{\mu\mu} \right). \quad (37)$$

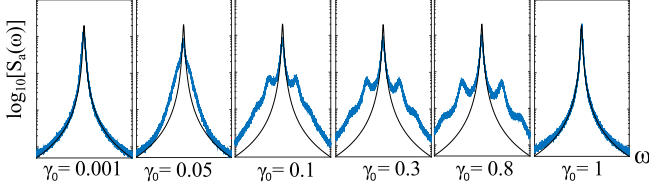


FIG. 3. (Color online) Simulated spectrum of the time-delayed model (33) with $\text{Re}C = 10$ and $R = 0.1$ (in arbitrary frequency units) at six values of γ_0 (using a base 10 logarithmic scale for the y axis). The noisy signal is the simulation result and the black curves are Lorentzian lineshapes with widths given by (29).

Therefore, the generalized α factor (which is responsible for linewidth enhancement due to coupling of amplitude and phase fluctuations) is given by

$$\tilde{\alpha}_\mu \equiv \frac{1}{R_{\mu\mu}} \left[\mathbb{B} \mathbb{A}^{-1} \mathbb{R} (\mathbb{B} \mathbb{A}^{-1})^T \right]_{\mu\mu}. \quad (38)$$

In the single-mode case ($M = 1$), this matrix formula reduces to the single-mode linewidth of the instantaneous model: $\frac{R}{2a_0^2} (1 + (\frac{B}{A})^2)$ [(29,30) in Sec. V.A].

The linewidth in the time-delayed (class B) model is precisely the same (neglecting side peaks) as in the instantaneous (class A) model. While this result was derived for single-mode class B semiconductor lasers using a phenomenological rate-equation framework [76], we prove that this is generally the case in the multimode inhomogeneous regime. Naively, one might expect to obtain different linewidths due to the longer time over which the fluctuations can grow. However, in appendix B we obtain a linewidth expression which is independent of the relaxation-oscillation dynamics, which demonstrates that there is a cancellation of two competing processes: as γ_{\parallel} decreases, amplitude fluctuations grow, but they are also averaged over longer periods of time so that their effect is smaller.

Fig. 3 presents the simulated spectrum of the time-delayed model in the homogeneous- γ limit, which is obtained by numerically integrating (33) (by applying a stochastic Euler scheme, as in Fig. 2). The width of the central peak of the spectrum matches our prediction (29), independent of the value of γ_0 . At intermediate relaxation rates, we also observe side peaks in the spectrum due to amplitude relaxation oscillations (RO), in addition to the central peak.

C. Side peaks in the time-delayed model

In class B lasers, amplitude fluctuations relax to steady state via relaxation oscillations [45] and, consequently, give rise to side peaks in the spectrum, in analogy with amplitude modulation of harmonic signals. Mathematically, the oscillation arises from the second-order ODE

generated by coupling of the δ and $\dot{\xi}$ equations (33,34), producing the coupled amplitude/gain oscillations. Using the same methods that we applied to calculate the linewidth of the central resonance peak (37), we also calculated the full side-peak spectrum in the multimode regime. Our formula is derived under the fairly general assumption that the central resonance peaks are narrower than the side peaks, which is the relevant regime for many lasers [76]. Although the derivation uses the same techniques as in appendix B, it is fairly involved and will be provided in a subsequent manuscript [84]; we only summarize here.

As was shown in Sec. III, far above threshold, the atomic relaxation rate (13) is enhanced and can even be dominated by the electromagnetic field. This modified relaxation rate, and in particular its spatial dependence due to hole-burning effects, has important implications on the RO spectrum which, to our knowledge, have not been treated before. For simplicity, we focus here on the case of $\alpha = 0$. (Note that α factor effects on the RO spectrum have been observed and analyzed using a phenomenological *homogeneous* time-delayed model in [76].)

In order to see how one can obtain a closed-form expression for the RO spectrum, recall that when calculating the spectrum of the central resonance peak in Sec. V.A, we neglected direct amplitude-fluctuation contributions in (26), i.e., in passing from the first to second step, we omitted a term of the form

$$\langle \delta(t) \delta(0) \rangle \cdot \left\langle e^{-i(\phi(t) - \phi(0))} \right\rangle. \quad (39)$$

Adding this term in (26), one finds that the full spectrum consists of an additional term, which is given by the convolution of the real-amplitude fluctuation spectrum $\langle \delta(t) \delta(0) \rangle$ and the spectrum of the central resonance peak. In the instantaneous model, the amplitude autocorrelation function $\langle \delta(t) \delta(0) \rangle$ decays exponentially in time [see (27)] and the omitted term results in near-constant background noise. However, in the time-delayed model, this neglected term is responsible for the RO side peaks.

For simplicity, consider first a model which can be solved straightforwardly; the single-mode homogeneous- γ time-delayed model [i.e., $\gamma(\mathbf{x}) \approx \gamma_0$ and $\int dx c(\mathbf{x}) = C$ as in (33)], which describes uniformly pumped single-mode lasers near threshold. Following the discussion in Sec. V.B, we can rewrite (33,34) as a set of linear equations and solve for $\delta(t)$, obtaining

$$\begin{aligned} \delta(t) = & \int dt' e^{-\frac{\gamma_0}{2}(t-t')} \\ & \times \left[\cosh \left(\frac{\Delta}{2}(t-t') \right) + \frac{\gamma}{\Delta} \sinh \left(\frac{\Delta}{2}(t-t') \right) \right] f_R(t'), \end{aligned} \quad (40)$$

where $\Delta \equiv \sqrt{\gamma_0^2 - 4A\gamma_0}$. In the limit of well-resolved side peaks (e.g., $\frac{R}{a_0^2} \ll \gamma_0 \ll A$), the amplitude autocor-

relation function is approximately

$$\langle \delta(t) \delta(0) \rangle \approx \frac{R}{2} \left[\frac{\sin \sqrt{\gamma_0 A} t}{\sqrt{\gamma_0 A}} + \frac{\cos \sqrt{\gamma_0 A} t}{\gamma_0} \right] \times e^{-\frac{\gamma_0 t}{2}}. \quad (41)$$

Thus, additional peaks in the spectrum arise at frequencies $\omega_{\text{RO}} = \omega_0 \pm \sqrt{A\gamma_0}$ with widths γ_0 . In the high-Q limit near threshold, A is proportional to the cavity decay rate κ , giving the expected behavior for the RO frequency. The side-peak amplitudes $\frac{R}{4} \left[\frac{1}{\sqrt{\gamma_0 A}} + \frac{1}{\gamma_0} \right]$ diverge in the limit of $\gamma_0 \rightarrow 0$ (that is, when amplitude fluctuations are not small compared to the steady-state mode amplitude), but this is also the regime in which our analysis of the spectrum (Sec. V.A-B) breaks down. The inset in Fig. 4b shows the simulated spectrum of the homogeneous time-delayed model (33) (the same data was also shown in Fig. 3, but we include here the theoretical formula for the side-peak spectrum). The exact numerical solution of (33) (blue curve) reproduces the analytic spectrum prediction of (42) (red curve).

In the limits of extremely small/large relaxation rates γ_0 (compared to A), the side peaks disappear. In the former limit, they merge with the central resonance peak and in the latter case, they merge with the background noise. This behavior can be explained by inspection of the δ and ξ equations (33,34) in the appropriate limits. When the relaxation rate is very large, the time-delayed model reduces to the instantaneous model, which represents the case where the atomic population follows the field adiabatically. In the opposite limit of extremely small relaxation, the field follows the atomic population adiabatically. In other words, a clear separation of atomic and optical time scales will result in the absence of RO side peaks.

In the most general spatially inhomogeneous time-delayed model, the full spectrum takes the simple form

$$S_a(\omega) = \frac{\Gamma}{\omega^2 + \left(\frac{\Gamma}{2}\right)^2} + \frac{\Gamma}{\omega^2 \left[1 - \int dx \frac{A(\mathbf{x})\gamma(\mathbf{x})}{\omega^2 + \left(\frac{\Gamma}{2} + \gamma(\mathbf{x})\right)^2} \right]^2 + \left[\int dx \frac{A(\mathbf{x})\gamma(\mathbf{x}) \left(\gamma(\mathbf{x}) + \frac{\Gamma}{2}\right)}{\omega^2 + \left(\gamma(\mathbf{x}) + \frac{\Gamma}{2}\right)^2} \right]^2} \quad (42)$$

where $A(\mathbf{x})$ is the real part of the local nonlinear coupling [defined in (23)], $\gamma(\mathbf{x})$ is the effective decay rate, and Γ is the central peak linewidth. (This formula is valid when the central resonance peak is narrower than the side peaks $\Gamma \ll \gamma_{\parallel}$.) Like our linewidth formula, this formula is easy to evaluate via spatial integrals of the SALT solutions.

While the homogeneous time-delayed model near threshold agrees with standard results on relaxation oscillations [76], the full model above threshold, combined with SALT, is able to include effects not contained in other treatments. As the pump is increased far above threshold, the effects of stimulated emission strongly increase the atomic relaxation rate, and spatial hole burn-

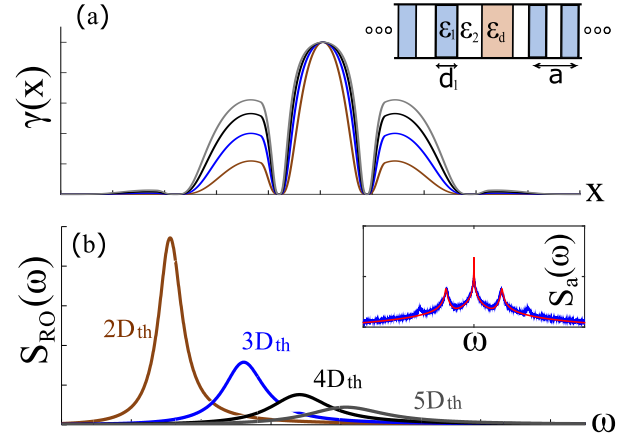


FIG. 4. (Color online) Dressed decay rate and the RO spectrum based on SALT solutions of a 1d PhC laser. Inset: a quarter-wave PhC (period $a = 1$ mm and alternating layers with permittivities $\epsilon_1 = 16 + 0.1i$ and $\epsilon_2 = 2 + 0.1i$ and thicknesses $d_1 = \frac{a\sqrt{\epsilon_2}}{\sqrt{\epsilon_1} + \sqrt{\epsilon_2}}$ and $d_2 = a - d_1$). The center region has permittivity $\epsilon_d = 3 + 0.1i$ and contains gain atoms with bandwidth $\gamma_{\perp} = 3 \text{ mm}^{-1}$ and resonance frequency $\omega_a = 25 \text{ mm}^{-1}$. (a) Dressed decay $\gamma(\mathbf{x})$ evaluated using (13) at five pump values ($2D_{\text{th}}$ brown, $3D_{\text{th}}$ blue, $4D_{\text{th}}$ black, and $5D_{\text{th}}$ gray). (b) Side-peak spectrum $S_{\text{RO}}(\omega)$ evaluated using (42) for the five pump values of (a). Inset: full simulated spectrum $S_a(\omega)$ on a semi-log scale (of base 10) of the homogeneous time-delayed model (33) with $\gamma_0 = 0.09$, $A = 10$, $B = 0$, $R = 0.01$ (in arbitrary frequency units). The noisy signal is the simulation result and the red curve is the theoretical lineshape (42).

ing causes that rate $\gamma(\mathbf{x})$ to vary substantially in space [see (13)]. These two effects cause both a shift and a broadening of the side peaks compared to the near-threshold result. Fig. 4 shows the dressed decay rate $\gamma(\mathbf{x})$ and the side-peak spectrum $S_{\text{RO}}(\omega)$ [as given by the second term of (42)], based on a SALT calculation of a one-dimensional photonic crystal (PhC) laser, at four different pump values well above threshold. [The pump value is controlled via the parameter D_p in (A3), and we denote the threshold value of D_p by D_{th} .] This type of cavity (depicted in the inset of Fig. 4a) supports a single mode at the simulated parameter regime, which is localized near the defect region. (Further discussion of this structure is given in Sec. VI.A below.) As can be seen from Fig. 4a, the decay rate $\gamma(\mathbf{x})$ is enhanced at high intensity regions (i.e., near the defect), and it increases further as the pump increases. Fig. 4b demonstrates the shifting and broadening of the side peaks.

VI. THE GENERALIZED α FACTOR

Our TCMT derivation of the linewidth formula yields a generalized α factor (38) which depends on the eigenmodes $\mathbf{E}_{\mu}(\mathbf{x})$ and eigenfrequencies ω_{μ} of the full non-

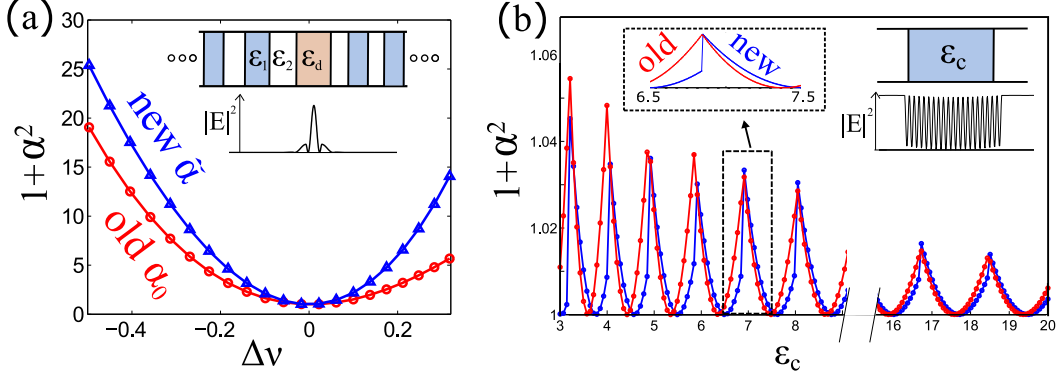


FIG. 5. (Color online) (a) The generalized (blue) and traditional (red) α factors of a PhC laser vs. relative detuning $\Delta\nu \equiv \frac{\omega_0 - \omega_a}{\omega_0}$. Upper inset: quarter-wave PhC geometry (see caption of Fig. 4). The gain parameters are $\gamma_\perp = 3 \text{ mm}^{-1}$ and a varying ω_a . Lower inset: intensity distribution of the lasing mode. (b) α factor for an open cavity laser vs. passive permittivity ϵ_c . Blue (red): generalized (traditional) α factor. Upper inset: dielectric slab, of permittivity ϵ_c , bounded by air on both sides, containing gain atoms with $\omega_a = 15 \text{ mm}^{-1}$, $\gamma_\perp = 3 \text{ mm}^{-1}$. Lower inset: intensity distribution of the lasing mode. Leftmost inset: enlarged segment of the main plot, around $\epsilon_c = 7$.

linear SALT equations. This is an advance over previous linewidth formulas; the *ab-initio* scattering-matrix linewidth formulas did not obtain an α factor [13, 14], whereas other traditional laser theories that derived α factors could not handle the full nonlinear equations [50]. Therefore, in the following section, we focus on the generalized α factor. We compare the generalized and traditional factors in Sec. VI.A, and then we evaluate these factors in the single-mode (Sec. VI.B) and multimode (Sec. VI.C) regimes.

A. Comparison with traditional α factor

Linewidth broadening due to amplitude-phase coupling (that is, the α factor linewidth enhancement) was first studied in the 1960s by Lax in the context of single-mode detuned gas lasers [9]. The Lax α factor is $1 + \alpha_0^2$, where α_0 is the normalized detuning of the lasing frequency from the atomic resonance, i.e., $\alpha_0 = \frac{\omega_0 - \omega_a}{\gamma_\perp}$, which is equal to the ratio of the real part of the gain permittivity to its imaginary part, or equivalently the ratio $\frac{\text{Re } \Delta n_g}{\text{Im } \Delta n_g}$, where Δn_g is the refractive index change due to fluctuations in the gain. Two decades later, Henry derived an amplitude-phase coupling enhancement factor of the same general type in semiconductor lasers [11], $\alpha_0 = \frac{\text{Re } \Delta n_g}{\text{Im } \Delta n_g}$, but in the latter case these refractive-index changes arise from carrier-density fluctuations and take a different form. Here, we are considering atomic gain media, so our α factor generalizes the Lax form.

The difference between our single-mode generalized α factor (30) and that of Lax arises because we take into account spatial variation in the gain permittivity due to spatial hole-burning and also the non-Hermitian (complex) nature of the lasing mode. Hence we expect our factor to reduce to the Lax factor in some limits. For

instance, consider the situation that was discussed in the last paragraph of Sec. II of a low-loss 1d Fabry-Pérot cavity laser, operating near threshold. In this case, the nonlinear coupling coefficient is approximately $C \approx \frac{-i\omega_0}{2\epsilon} \frac{\int \frac{\partial \epsilon}{\partial |a|^2} \mathbf{E}_0^2}{\int \mathbf{E}_0^2}$, and one can show that the generalized α factor is $\tilde{\alpha} = \frac{\text{Im } C}{\text{Re } C} \approx \frac{\text{Re } \Delta \epsilon}{\text{Im } \Delta \epsilon}$ (the last approximation is valid since in essentially all realistic cavities, the modes can be chosen to be predominantly real, i.e., have small imaginary parts).

In many cases, however, our $\tilde{\alpha}$ deviates from the traditional factor α_0 . An obvious example is when the lasing frequency precisely coincides with the atomic resonance frequency. In this case, the traditional factor vanishes, but $\tilde{\alpha}$ does not necessarily vanish. In the next section, we calculate and discuss the characteristic properties of the generalized α factor for two 1d laser structures.

B. Generalized single-mode α factor

In this section, we evaluate the differences between the generalized and traditional α factors in 1d model systems. We solve the full nonlinear SALT equations using our recent finite-difference frequency-domain (FDFD) SALT solver [21].

The generalized factor $\tilde{\alpha}$ can deviate significantly from the traditional factor α_0 when the latter is large (a similar argument was made in [44]). To see this, let us write the nonlinear coupling coefficient qualitatively as $C \propto (1 + i\alpha_0)(1 + i\beta)$, where the term $1 + i\alpha_0$ is associated with the atomic lineshape $\frac{\gamma_\perp}{\omega_0 - \omega_a + i\gamma_\perp}$, and the term $1 + i\beta$ is a complex factor due to the remaining integral factors (we refer to the latter term as the *modal contribution* to the α factor). Typically $\beta \ll 1$ and, consequently, the generalized factor is approximately $\tilde{\alpha} \approx \alpha_0 + \beta(1 + \alpha_0^2)$, so the difference between the generalized and traditional

factors grows quadratically with α_0 .

To verify this argument, we study a model system in which the magnitude of α_0 can be controlled. Consider a quarter-wave dielectric photonic crystal (PhC), with a defect at the center of the structure (the geometry is depicted in the upper inset of Fig. 5a, similar to the structure that was studied in Fig. 4). Adding enough layers of the periodic structure on each side of the defect to mimic an infinite structure, one finds that the system has a localized mode in the vicinity of the defect (lower inset), whose resonance frequency is fixed to a real value within the energy gap [33]. To study finite-threshold lasers, we introduce gain and some passive loss (i.e., a positive imaginary permittivity term, which pushes the resonance poles away from the real axis in the complex plane). Since the resonance frequency of the defect mode is fixed by the geometry, by varying the resonance frequency of the gain, we control the detuning of the lasing mode from the atomic resonance, thus controlling the size of α_0 . As demonstrated in the figure, the deviation $|\tilde{\alpha} - \alpha_0|$ grows as the detuning $\Delta\nu \equiv \frac{\omega_0 - \omega_a}{\omega_0}$ increases.

The openness of the cavity also results in an enhancement of the α factor; the more open it is, the larger is the necessary imaginary part of the lasing mode, which causes a deviation from the standard formula. In order to test this prediction, we evaluate the generalized α factor for an open-cavity laser (Fig. 5b), where we can control the radiative loss rate through the cavity walls and, consequently, this part of the modal contribution to $\tilde{\alpha}$. We consider a cavity which consists of a dielectric slab (with permittivity ε_c) surrounded by air on both sides, with gain spread homogeneously inside the slab (upper right-most inset). The reflectivity of the cavity walls is determined by the difference in cavity and air permittivities $\Delta\varepsilon = \varepsilon_c - \varepsilon_0$. For relatively small dielectric mismatch, the cavity is relatively low- Q and our α factor differs significantly from the Lax factor. As $\Delta\varepsilon$ increases and the cavity Q increases, the generalized α factor converges to the original factor, so that the red and blue curves in the figure overlap.

Unlike a photonic-crystal defect-mode cavity where there is a finite bandwidth of confinement [33], this dielectric cavity has an infinite number of possible lasing resonances and thus when we sweep $\Delta\varepsilon$, the α factor peaks periodically. This is because the free spectral range of the cavity is $\Delta\omega \approx \frac{2\pi}{\sqrt{\varepsilon_c}L}$ [85] and, therefore, changing ε_c corresponds to shifting the passive resonances and, consequently, the lasing modes. Every time a lasing mode crosses an atomic resonance, α_0 vanishes and correspondingly $\tilde{\alpha}$ becomes very small. The traditional factor is maximized when the atomic resonance is equidistant from two passive modes. The peak value is proportional to the free spectral range and, therefore, we find that it is proportional to $1/\sqrt{\varepsilon_c}$. This type of effect may not have been observed previously because in macroscopic cavities, the cavity resonances are very dense on the scale of the gain bandwidth, so the lasing mode can never be substantially detuned. However, in microcavities with

large free spectral range, this could be an important effect. Another intriguing property of the generalized α factor is that it varies discontinuously at the peaks (as is shown more clearly in the upper left-most inset). The traditional factor α_0 depends only on the mode detuning from resonance, so it approaches the same value on different sides of the peak. In contrast, the generalized factor $\tilde{\alpha}$ depends on the mode profile \mathbf{E}_μ , which differs between the two interchanging laser modes on different sides of the peak, producing the observed asymmetry.

C. Generalized multimode α factor

Our *multimode* linewidth formula includes linewidth corrections from neighboring modes, which enter through the generalized α factor (since phase fluctuations in each of the modes couple to amplitude fluctuations in all other modes due to saturation of the gain). According to the traditional ST formula (2), when phase cross-correlations between different modes are neglected, each resonance-peak width is inversely proportional to the corresponding modal output power. We find that when phase cross-correlations are included, the linewidth of each mode is a sum of inverse output powers of all the other modes—a type of multimode Schawlow–Townes relation. To see how this comes about, recall that the generalized α factor, as given by (38), is proportional to $[\mathbb{B}\mathbb{A}^{-1}\mathbb{R}(\mathbb{B}\mathbb{A}^{-1})^T]_{ii}$. We show in appendix C that individual factors in the product scale as $[\mathbb{B}\mathbb{A}^{-1}]_{ij} \propto \frac{a_{i0}}{a_{j0}}$, where a_{j0} is the steady-state amplitude of the j 'th mode. Therefore, the multimode α factor is proportional to the sum: $a_{i0}^2 \sum_j \frac{(\text{const}) \times R_{jj}}{a_{j0}^2}$, i.e., a sum over terms which scale as inverse output powers.

In the two-mode case, the linewidth formula for a lasing mode in the presence of a neighboring mode is given explicitly by

$$\Gamma_1 = \frac{R_{11}}{2a_{10}^2} + \frac{R_{11}}{2a_{10}^2} \left[\frac{C_{11}^I C_{22}^R - C_{21}^I C_{21}^R}{C_{11}^R C_{22}^R - C_{12}^R C_{21}^R} \right]^2 + \frac{R_{22}}{2a_{20}^2} \left[\frac{C_{11}^R C_{12}^I - C_{11}^I C_{12}^R}{C_{11}^R C_{22}^R - C_{12}^R C_{21}^R} \right]^2, \quad (43)$$

where $C_{ij}^R \equiv \text{Re } C_{ij}$ and $C_{ij}^I \equiv \text{Im } C_{ij}$. (A similar expression was derived in [39], by using a phenomenological version of the two-mode TCMT equations.) As predicted by the multimode ST relation, the last term in (43) is inversely proportional to the output power of the second mode a_{20}^2 . This term becomes significant when the power in the first mode greatly exceeds the power in the second mode (i.e., when $P_1 \gg P_2$), correcting the unrealistic Schawlow–Townes prediction that the linewidth vanishes when $P_1 \rightarrow \infty$; a similar argument was made in [39]. Fig. 6a presents the spectrum of a two-mode instantaneous model (15) in the parameter regime where

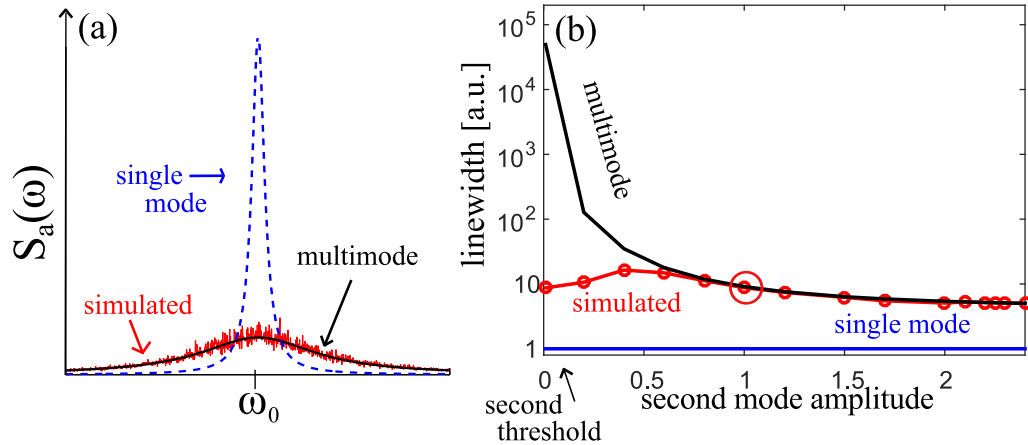


FIG. 6. (Color online) (a) Spectrum near a resonance peak in the presence of an additional mode. Numerical simulations of (15) (red curve) and analytic single-mode (blue) and multimode (black) formulas. The simulation parameters are chosen so that there are two lasing modes with the same steady-state amplitudes $a_{k0} = 1$ and diffusion coefficients $R_{kk} = 0.05$, and with substantial cross correlations: $C_{kk} = 5$, $C_{kl} = 4 + 4i$, $k \neq l$ (in arbitrary frequency units). (b) Linewidth of central resonance peak vs. output power in the neighboring mode [$a_{10} = 1$ and $a_{20} \in (0, 3]$]. Simulated spectrum (red) and analytic single-mode and multimode formulas (blue and black curves). The point $a_{10} = a_{20} = 1$ is encircled, and corresponds to the parameter values of Fig. 6a.

cross-correlations between the two modes are significant. The linewidth of the simulated spectrum (red curve) is in complete agreement with the generalized formula (43) (black curve), but deviates substantially from the single-mode formula (29) (blue curve). In order to reach the regime where this deviation is substantial, in practice, one needs to design a cavity in which the two lasing modes have comparable amplitudes and detunings from the atomic resonance frequency.

Eq. (43) predicts an unphysical divergence near the second threshold, i.e., when $a_{20} \rightarrow 0$ (see black curve in Fig. 6b). In retrospect, this singularity is to be expected, since the assumptions of our derivation break down in this limit. (Note that an equivalent divergence was present in [39].) In calculating the phase variance, we assumed that amplitude fluctuations in all modes were small compared to the steady-state amplitudes ($\delta_I \ll a_{i0}$), and this assumption is no longer valid near threshold. The N-SALT TCMT equations (11) are still valid, however—it is only their analytical solution for $\langle \Phi \Phi^T \rangle$ that is problematic. Therefore, we study the threshold regime numerically, via stochastic simulations of the N-SALT TCMT equations. As shown in Fig. 6b, the simulated linewidth of the first mode approaches a finite value near the second threshold (red curve), and this value is significantly larger than the linewidth prediction one obtains when neglecting the second mode (blue curve). Even at the threshold, noise in the second mode mixes with the first mode through off-diagonal nonlinear coupling terms, thus increasing the linewidth.

Linewidth enhancement at the thresholds of neighboring lasing-modes suggests that the linewidth must also be enhanced *below* the modal thresholds [in the regime

where radiation from non-lasing modes is incoherent, commonly called amplified spontaneous emission (ASE)]. We believe that this phenomenon could be explored using a future generalization of our formalism, with some modifications (extending earlier work [28, 40] on linewidth enhancement from ASE).

VII. FULL-VECTOR 3D EXAMPLE

In order to illustrate the full generality of our approach, we apply it in this section to study a three-dimensional photonic-crystal (PhC) laser. The steady-state properties of this system (i.e., the lasing threshold and mode characteristics) were previously explored in [21]. We use those solutions here to calculate the laser linewidth [using (3)], and we compare the relative contributions of the various correction factors.

The simulated PhC consists of a dielectric slab patterned by a hexagonal lattice of air holes (Fig. 7a). A defect is introduced by decreasing the radii of seven holes at the center of the structure [86], giving rise to a doubly-degenerate mode which is situated at the defect (spatially) and in the bandgap of the lattice (spectrally). We select the TE-like mode out of the degenerate pair by imposing even and odd reflection symmetry at $x = 0$ and $y = 0$ respectively, as well as an even reflection symmetry at $z = 0$. Staying close to a potential experimental realization, we choose the pump profile to be uniform inside the high-index dielectric near the defect region, and zero elsewhere. We solve the SALT equations using our scalable FDFD solver, and track the evolution of the first lasing mode upon increasing the pump strength from zero

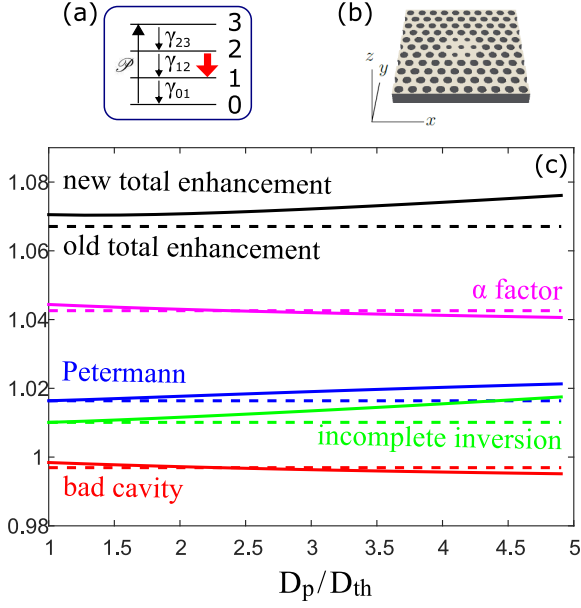


FIG. 7. (Color online) Linewidth correction factors for a 3d PhC laser. (a) The PhC consists of a hexagonal lattice of air holes (with period $a = 1$ mm and radius 0.3 mm) in a dielectric medium with index $n = \sqrt{\epsilon_c} = 3.4$. The slab has a thickness of 0.5 mm, with air above and below, terminated by PML absorbers. A cavity is formed by seven holes of radius of 0.2 mm. The pump is non-zero in the hexagonal region for height 2mm in the z -direction. (Borrowed from [21].) (b) Schematics of a 4-level gain medium. Levels $|1\rangle$ and $|2\rangle$ form the lasing transition, with resonance frequency $\omega_a = 1.5$ mm $^{-1}$ and polarization decay $\gamma_{\perp} = 2.0$ mm $^{-1}$. The population decay rates are $\gamma_{01}/\gamma_{12} = \gamma_{23}/\gamma_{12} = 10^2$ and $\gamma_{12}/\gamma_{\perp} = 10^{-2}$. The pump rate \mathcal{P} is varied in the range $\mathcal{P}/\gamma_{12} = 0.4 \dots 2.2$. (c) Traditional (dashed) and generalized (solid) correction factors, as defined in Table. I. The total correction is defined as the product of the (traditional and generalized) Petermann, α , bad-cavity, and incomplete-inversion factors. The x axis is the D_p/D_{th} , where D_p is the SALT effective pump parameter (see text) and D_{th} is the effective threshold pump.

to five times the first-threshold value.

Typically, realistic laser structures do not use 2-level gain media, but employ a more complex optical scheme which involves multiple levels and transitions in order to achieve significant inversion and depletion of the ground-state population. In this section, we apply our formalism to a 4-level gain medium (Fig. 7b), using a generalization of SALT [73], which finds the stationary multimode lasing properties of an N -level gain medium. As shown in [73], an N -level system can be mapped into an effective 2-level system, which obeys the (2-level) SALT equations with renormalized pump (D_p) and atomic relaxation rates (γ_{\parallel}). Consequently, the linewidth of a 4-level laser will be given by our generalized formula (3) with the appropriately renormalized coefficients. By choosing the decay rate between the lasing transition levels (γ_{12} in Fig. 7b) to be much smaller than the decay rates into

the upper (γ_{23}) and out of the lower (γ_{01}) states, we can achieve substantial inversion and ground-state depletion. Consequently, the incomplete-inversion factor is approximately $n_{sp} \approx 1$, close to typical measured values [47].

Fig. 7c presents the traditional and new correction factors (dashed and solid lines respectively), as defined in Table. I. We find that those factors are relatively small for this system and, consequently, the deviations between the new and traditional factors are small. A small Petermann factor arises since the first lasing mode has a relatively high quality factor (i.e., the cold-cavity resonance pole is at $\omega_0 = 1.725 - 0.00512i$ mm $^{-1}$ with a quality factor of $Q \approx 700$, in agreement with experimental realization [86]). Moreover, the cold-cavity resonance lies well within the gain bandwidth, resulting in small α and bad-cavity corrections. The generalized factor $\tilde{\alpha}$ (solid purple line) is obtained from from (17,30). Deviations of $\tilde{\alpha}$ from the traditional factor $\alpha_0 \equiv \frac{|\omega_0 - \omega_a|}{\gamma_{\perp}}$ (dashed purple line) are due to modal contributions to the α factor (see Sec. VI.B). The generalized Petermann factor (full blue curve) is compared against the traditional factor (dashed blue line), which is expressed in terms of the SALT mode (instead of the passive cavity mode). The cavity region is taken to be the entire high-index medium. (Note the the generalized and traditional factors agree at threshold). Both the Petermann and α factors increase the linewidth. However, the generalized and traditional bad-cavity factors (full and dashed red curves respectively) lead to linewidth reduction.

Last, we evaluate the incomplete-inversion factor \tilde{n}_{sp} . The inversion $D(x)$ is found from the SALT solutions of the effective 2-level system. The excited state population $N_2(x)$ can be derived straightforwardly, using the results of [73] as follows. Assuming that the populations in the non-lasing levels $|0\rangle$ and $|3\rangle$ are at steady-state, one can express those populations in terms of the populations in the lasing transition $|1\rangle$ and $|2\rangle$. Then, by invoking the density conservation condition, $\sum_i N_i = n$, where n is the atom number density and N_i are the individual level populations, one finds that the population in $|2\rangle$ is given by

$$N_2 = \frac{n + \tau D}{1 + \tau}, \quad (44)$$

where $\tau \equiv 1 + \frac{2\gamma_{01}}{\gamma_{23}} + \frac{\gamma_{01}}{\mathcal{P}}$. Having obtained expressions for D and for N_2 , we have all that is needed to calculate the incomplete-inversion factor \tilde{n}_{sp} . We define the “linear incomplete-inversion factor” (n_{sp} dashed green line) as the ratio $\frac{N_2(D_p)}{D_p}$, [i.e., both the excited-state population (44) and the inversion are evaluated at $D = D_p$, neglecting hole-burning effects]. The “nonlinear incomplete-inversion factor” (\tilde{n}_{sp} solid green line) is defined in Table. I. The nonlinear factor \tilde{n}_{sp} coincides with the linear factor n_{sp} at threshold, but exceeds the traditional factor at higher pumps. We also plot the total linewidth correction, which is defined as the product of the (traditional and new) Petermann, α , bad-cavity, and

incomplete-inversion factors.

VIII. CONCLUDING REMARKS

We presented a generalized multimode linewidth formula, obtained from the N-SALT TCMT equations for the lasing mode amplitudes, which we derived starting from the Maxwell–Bloch equations and using the fluctuation–dissipation theorem to determine the statistical properties of the noise. Our generalized linewidth formula (3) reduces to the traditional formula (2) for low-loss cavities and simple lasing structures, but deviates significantly from the traditional theories for high-loss wavelength-scale laser cavities. By basing our derivation on the SALT steady-state lasing modes, it is possible to apply our formula to cavities of arbitrarily complex geometry (e.g., photonic crystal or microdisk lasers [15–18]) and arbitrary openness (e.g. random lasers [43]). Also, since SALT includes to high accuracy the effects of spatial hole-burning, our formula includes both gain saturation and the spatial variation of the gain permittivity well above threshold, plus all effects due to modal couplings. From a computational point of view it is important to point out that our formula is analytical and can be evaluated immediately from the output of a numerical SALT calculation without any significant computational effort. A manuscript describing a brute-force numerical validation of our theory against numerical solution of the Maxwell–Bloch equations is currently being prepared [41]. Given only the laser geometry, the pumping profile, and characteristic properties of the gain (i.e., its resonance frequency ω_a and decay rate γ_\perp), our formula enables linewidth calculation, including a generalized α -factor and accounting for temperature variations, at a level of generality that was not possible before. This generality is most important, of course, in cases where the new result is substantially different than previous theories, and it would be interesting to study laser cavities in which the discrepancy is as large as possible.

One such case is that of lasers which contain exceptional points (EPs) in their spectrum, which are points of degeneracy where two (or more) eigenfrequencies and eigenfunctions coalesce [65, 87]. EPs in laser systems have been explored recently, both theoretically [88] and experimentally [89]. At the EP, the modes become self-orthogonal and that causes the denominator of (3) to vanish and is already known to greatly enhance the Petermann factor [90]. Since a similar denominator appears in the integrals defining our generalized α factor (12,30), we expect that our $\tilde{\alpha}$ will differ substantially from previous results near an EP (and similarly for the inhomogeneous-temperature correction).

An important and exciting addition to the theory would be a treatment of amplified spontaneous emission (ASE) from modes below threshold; we believe this can be achieved by deriving TCMT equations for below-threshold (passive) modes, in which there is no steady-

state oscillation (generalizing previous ASE work which used simplified models [28, 40]). Incorporating the ASE contribution to the spectrum will allow us to follow the noise through the lasing thresholds, correcting the unphysical divergence which was discussed in Sec. VI.B. More importantly, treating below threshold ASE should allow an ab-initio theory of LEDs in arbitrary cavities.

Future work could also incorporate several additional corrections that were not treated in this paper. Our derivation applies to isotropic materials described by a scalar permittivity ε , but extension to anisotropic permittivity $\hat{\varepsilon}$, magnetic permeability ($\hat{\mu}$), and even bianisotropic materials would be very straightforward (e.g., for an anisotropic $\hat{\varepsilon}$, the only change is that $\varepsilon \mathbf{E}^2$ factors and similar are replaced by $\mathbf{E} \cdot (\hat{\varepsilon} \mathbf{E})$ etcetera, as in [55]). As discussed in Sec. V.C, we are also able to exploit our framework to analytically solve for the relaxation-oscillation side-peak spectra, and are currently preparing a manuscript presenting this analysis [84]. We believe it will be possible to extend our formalism to handle non-Lorentzian lineshapes arising from frequency dependence (correlations) in the noise within the laser linewidth [60–64], as also discussed in Sec. IV. Instead of treating the noise spectrum $S_{\mathbf{F}}(\omega)$ as a constant $S_{\mathbf{F}}(\omega_\mu)$, one needs to include a first-order correction, e.g., by Taylor expanding $S_{\mathbf{F}}(\omega)$ around ω_μ ; it might be convenient to fit $S_{\mathbf{F}}(\omega)$ to a Lorentzian matching the amplitude and slope at ω_μ , since the Fourier transform of a Lorentzian is an exponential that should be easy to integrate. Finally, as noted above, although our derivation was for the two-level Maxwell–Bloch equations, a similar approach should apply to more complex gain media (including multi-level atoms [73], multiple lasing transitions, and gain diffusion [74].) The N-SALT linewidth theory can be generalized to account for these laser models following along the lines of our approach here.

ACKNOWLEDGMENTS

This work was partially supported by the Army Research Office through the Institute for Soldier Nanotechnologies under Contract No. W911NF-13-D-0001. ADS and AC acknowledge the support of NSF Grant No. DMR-1307632. CYD acknowledges the support of Singapore NRF Grant No. NRFF2012-02. The authors would like to thank Bo Zhen, Aristeidis Karalis, Amir Rix, Owen Miller, and Homer Reid for helpful discussions.

APPENDIX A: DERIVATION OF N-SALT TCMT

In this appendix, we derive the TCMT equations for the lasing mode amplitudes. Our starting point is the Maxwell–Bloch equations [2, 22], which describe the dynamics of the electromagnetic field in a resonator interacting with a two-level gain medium:

$$\nabla \times \nabla \times \mathbf{E} + \varepsilon_c \ddot{\mathbf{E}} = -4\pi \ddot{\mathbf{P}} + \mathbf{F}_S, \quad (\text{A1})$$

$$\dot{\mathbf{P}} = -i(\omega_a - i\gamma_\perp)\mathbf{P} - \frac{i\gamma_\perp}{4\pi}\mathbf{E}D, \quad (\text{A2})$$

$$\dot{D} = -\gamma_\parallel [D_p - D + 2\pi i(\mathbf{E} \cdot \mathbf{P}^* - \mathbf{E}^* \cdot \mathbf{P})], \quad (\text{A3})$$

where \mathbf{E} is the electromagnetic field, while \mathbf{P} and D are the atomic polarization and population inversion. (From here on, for brevity, we refer to D as the “inversion.”) ω_a is the atomic resonance frequency, and γ_\perp and γ_\parallel are the population and inversion relaxation rates. D_p is the external pump, which determines the steady-state inversion, and ε_c is the passive dielectric permittivity. The field, polarization and inversion are measured in their natural units: $e_c = p_c = \hbar\sqrt{\gamma_\parallel\gamma_\perp}/(2g)$ and $d_c = \hbar\gamma_\perp/(4\pi g^2)$ respectively, where g is the atomic dipole matrix element [19–21]. We introduce spontaneous emission noise by including a random source term $\mathbf{F}_S = 4\pi\frac{\partial\mathbf{J}}{\partial t}$ in (A1), written in the frequency domain as

$$\hat{\mathbf{F}}_S(\mathbf{x}, \omega) = -i4\pi\omega\hat{\mathbf{J}}(\mathbf{x}, \omega), \quad (\text{A4})$$

where $\hat{\mathbf{J}}(\mathbf{x}, \omega)$ is a random fluctuating current, and the correlations of $\hat{\mathbf{F}}_S(\mathbf{x}, \omega)$ are given by the FDT.

Steady-state ab-initio laser theory (SALT) handles the noise-free regime of the Maxwell–Bloch equations (i.e., $\hat{\mathbf{F}}_S = 0$) and reduces this set of coupled equations to a frequency-domain nonlinear generalized eigenvalue problem for the electric field $\hat{\mathbf{E}}$ (as reviewed in Sec. 1.1). When noise is introduced ($\hat{\mathbf{F}}_S \neq 0$), the cavity field is perturbed from steady-state and the nonlinear permittivity is modified (Sec. 1.2). This gives rise to a restoring force (denoted $\hat{\mathbf{F}}_{\text{NL}}$), which we calculate in Sec. 1.3. The noise-driven field $\hat{\mathbf{E}}$ is then found by integrating the Green’s function (derived in Sec. 1.4) over the noise terms $\hat{\mathbf{F}}_S$ and $\hat{\mathbf{F}}_{\text{NL}}$. Finally, the TCMT equations are obtained by transforming back into the time domain (Sec. 1.5).

1. Review of SALT

We begin by reviewing the steady-state theory. In the SALT approach, the steady-state electromagnetic field is expressed as a superposition of a finite number of lasing modes:

$$\mathbf{E}_0(\mathbf{x}, t) = \sum_{\mu} \mathbf{E}_{\mu}(\mathbf{x}) a_{\mu 0} e^{-i\omega_{\mu} t}, \quad (\text{A5})$$

where $\mathbf{E}_0(\mathbf{x}, t)$ denotes the steady-state field and $a_{\mu 0}$ are the steady-state modal amplitudes. The lasing modes $\mathbf{E}_{\mu}(\mathbf{x})$ are real frequency solutions of the nonlinear eigenvalue problem

$$[\nabla \times \nabla \times - \omega_{\mu}^2 \hat{\varepsilon}_0(\omega_{\mu}, a_0)] \mathbf{E}_{\mu}(\mathbf{x}) = 0, \quad (\text{A6})$$

with outgoing boundary conditions. The effective permittivity has a linear (passive) term ε_c and a nonlinear (\mathbf{E} -dependent) gain term:

$$\hat{\varepsilon}_0(\omega, a_0) = \varepsilon_c + \frac{\gamma_\perp}{\omega - \omega_a + i\gamma_\perp} D_0(a_0). \quad (\text{A7})$$

The steady-state inversion $D_0(a_0)$ [which is a notation shortcut for $D_0(\{\mathbf{E}_{\mu}\}, \{\omega_{\mu}\}, \{a_{\mu 0}\})$] is given by

$$D_0(a_0) = \frac{D_p}{1 + \sum_{\mu} \frac{\gamma_\perp^2}{(\omega_{\mu} - \omega_a)^2 + \gamma_\perp^2} |a_{\mu 0}|^2 |\mathbf{E}_{\mu}|^2}. \quad (\text{A8})$$

To avoid possible confusion, note that in previous SALT works, the steady-state inversion was denoted by D and D_0 was the external pump parameter, whereas in this work, D_0 is the steady-state inversion and D_p is the external pump parameter.

2. Noise-driven Maxwell-Bloch equations

In the presence of a small noise source, the electric field and polarization can be written as superpositions of the steady-state lasing modes with time-dependent amplitudes $a_{\mu}(t)$ and $b_{\mu}(t)$:

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= \sum_{\mu} \mathbf{E}_{\mu}(\mathbf{x}) a_{\mu}(t) e^{-i\omega_{\mu} t} \\ \mathbf{P}(\mathbf{x}, t) &= \sum_{\mu} \mathbf{P}_{\mu}(\mathbf{x}) b_{\mu}(t) e^{-i\omega_{\mu} t}. \end{aligned} \quad (\text{A9})$$

Substituting the perturbation ansatz (A9) into the polarization equation (A2), we obtain

$$(\dot{b}_{\mu} + i\omega_{\mu} b_{\mu}) \mathbf{P}_{\mu} = -i(\omega_a - i\gamma_\perp) b_{\mu} \mathbf{P}_{\mu} - \frac{i\gamma_\perp a_{\mu}}{4\pi} \mathbf{E}_{\mu} D. \quad (\text{A10})$$

Taking the Fourier transform and rearranging terms, we find

$$\tilde{B}_{\mu} \mathbf{P}_{\mu} = \frac{1}{4\pi} \frac{\gamma_\perp}{\omega - \omega_a + i\gamma_\perp} \hat{a}_{\mu} * \hat{D} \mathbf{E}_{\mu}, \quad (\text{A11})$$

where we have introduced the shifted frequency $\omega \equiv \omega_{\mu} + \Omega$ and the Fourier-domain envelopes $\hat{a}_{\mu}(\Omega) = \hat{a}_{\mu}(\omega - \omega_{\mu})$, $\tilde{B}_{\mu}(\Omega)$ and $\hat{D}(\Omega)$. The asterisk $*$ denotes a convolution.

Next, consider Eq. (A1) in the frequency domain

$$\nabla \times \nabla \times \hat{\mathbf{E}} - \omega^2 \varepsilon_c (\hat{\mathbf{E}} + 4\pi \hat{\mathbf{P}}) = \hat{\mathbf{F}}_S. \quad (\text{A12})$$

When the spacing between adjacent lasing modes is much larger than their linewidths, a noise source with frequency $\omega \approx \omega_{\mu}$ excites only the mode $\mathbf{E}_{\mu}(\mathbf{x})$. Equivalently, the Green’s function can be approximated by the contribution of the single pole at ω_{μ} . (Note that we require only that the peaks in the laser spectrum above threshold are non-overlapping; we do not require isolated resonances in the passive cavity spectrum.) Therefore, at frequencies $\omega \approx \omega_{\mu}$, we can substitute (A11) into (A12) and obtain an effective equation for the noise-driven field $\hat{\mathbf{E}}_{\mu}(\mathbf{x}, \omega)$:

$$[\nabla \times \nabla \times - \omega^2 \varepsilon(\omega, a)] \hat{\mathbf{E}}_{\mu}(\mathbf{x}, \omega) = \hat{\mathbf{F}}_S(\mathbf{x}, \omega), \quad (\text{A13})$$

where the effective permittivity $\varepsilon(\omega, a)$ is given by

$$\varepsilon(\omega, a) \hat{\mathbf{E}}_\mu(\mathbf{x}, \omega) = \left[\varepsilon_c \hat{a}_\mu + \frac{\gamma_\perp}{\omega - \omega_a + i\gamma_\perp} \hat{D} * \hat{a}_\mu \right] \mathbf{E}_\mu(\mathbf{x}). \quad (\text{A14})$$

The second variable of $\varepsilon(\omega, a)$ denotes the implicit dependence of ε on the modal amplitudes a_μ through the Fourier transform of the inversion \hat{D} . We calculate \hat{D} explicitly in the next section.

3. The atomic inversion

The noise source $\hat{\mathbf{F}}_S$ perturbs the modal amplitudes a_ν from steady state, causing a change in the atomic inversion D . We neglect dispersion corrections to D (which amounts to setting $\dot{b}_\mu = 0$ in (A10) [37]) as these corrections do not affect the linewidth formula to leading order in the noise [see discussion following (9) in the main text]. From (A3) and (A10), we obtain

$$\dot{D} = -\gamma_\parallel \left[D - D_p + \sum_\nu \frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |a_\nu|^2 D |\mathbf{E}_\nu|^2 \right]. \quad (\text{A15})$$

In order to solve (A15), we linearize the time dependent products $|a_\nu|^2 D$ in the sum around the steady state $|a_\nu|^2 D \approx a_{\nu 0}^2 D_0 + D_0(|a_\nu|^2 - a_{\nu 0}^2) + a_0^2(D - D_0)$, where D_0 is the steady state (SALT) inversion (A8). To simplify the notation, we define the local decay rate

$$\gamma(\mathbf{x}) \equiv \gamma_\parallel \left(1 + \sum_\nu \frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |a_{\nu 0}|^2 |\mathbf{E}_\nu|^2 \right). \quad (\text{A16})$$

The second term in (A16) gives precisely the increased atomic decay rate due to stimulated emission. Using the definitions above, (A15) becomes

$$\begin{aligned} \dot{D} &= -\gamma(\mathbf{x})(D - D_0) \\ &\quad - \gamma_\parallel D_0 \cdot \sum_\nu \frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |\mathbf{E}_\nu|^2 (|a_\nu|^2 - |a_{\nu 0}|^2), \end{aligned} \quad (\text{A17})$$

which we can integrate, and obtain

$$\begin{aligned} D &= D_0 + \sum_\nu D_0 \left(\frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |\mathbf{E}_\nu|^2 \right) \times \\ &\quad \gamma_\parallel \int^t dt' e^{-\gamma(\mathbf{x})(t-t')} (|a_{\nu 0}|^2 - |a_\nu(t')|^2). \end{aligned} \quad (\text{A18})$$

Having derived an explicit expression for $D(t)$, we substitute its Fourier transform \hat{D} into the effective permittivity (A14) and obtain

$$\varepsilon(\omega, a) \hat{\mathbf{E}}_\mu \approx \varepsilon(\omega, a_0) \hat{\mathbf{E}}_\mu + \sum_\nu \chi_\nu(\omega, a_0) \widehat{\Delta a}_\nu * \hat{\mathbf{E}}_\mu, \quad (\text{A19})$$

where $\varepsilon(\omega, a_0)$ is the steady-state SALT permittivity which was defined in (A7), $\chi_\nu(\omega, a_0)$ is the permittivity differential due to deviation in the modal amplitude a_ν [which we denote by “ $\frac{\partial \varepsilon}{\partial |a|^2}$ ” in the text, e.g., in (12)]:

$$\chi_\nu \equiv \frac{\gamma_\perp}{\omega - \omega_a + i\gamma_\perp} D_0 \left(\frac{\gamma_\perp^2}{(\omega_\nu - \omega_a)^2 + \gamma_\perp^2} |\mathbf{E}_\nu|^2 \right) \frac{\gamma_\parallel}{\gamma(\mathbf{x})}, \quad (\text{A20})$$

and $\widehat{\Delta a}_\nu$ is the Fourier transform of the time-averaged modal deviation from steady state

$$\Delta a_\nu = \gamma(\mathbf{x}) \int^t dt' e^{-\gamma(\mathbf{x})(t-t')} (|a_{\nu 0}|^2 - |a_\nu(t')|^2). \quad (\text{A21})$$

Substituting the permittivity expansion (A19) into Maxwell's equation (A13), we obtain

$$[\nabla \times \nabla \times - \omega^2 \varepsilon(\omega, a_0)] \hat{\mathbf{E}}_\mu(\mathbf{x}, \omega) = \hat{\mathbf{F}}_{\text{NL}}(\mathbf{x}, \omega) + \hat{\mathbf{F}}_S(\mathbf{x}, \omega), \quad (\text{A22})$$

where the nonlinear restoring force is

$$\hat{\mathbf{F}}_{\text{NL}}(\mathbf{x}, \omega) = \omega^2 \sum_\nu \chi_\nu(\omega, a_0) \widehat{\Delta a}_\nu * \hat{\mathbf{E}}_\mu(\mathbf{x}, \omega). \quad (\text{A23})$$

The left-hand side of (A22) is just the linearized steady-state equation (A6), and the nonlinear correction to the effective permittivity due to the noise $\hat{\mathbf{F}}_S$ appears as an additional source term $\hat{\mathbf{F}}_{\text{NL}}$. As noted above, the noise-driven field $\hat{\mathbf{E}}_\mu$ is found by integrating the Green's function of the steady-state equation (A6) over the noise terms $\hat{\mathbf{F}}_{\text{NL}}$ and $\hat{\mathbf{F}}_S$. In the following section we derive an approximate formula for the Green's function.

4. The linearized steady state Green's function

The single-pole approximation of the Green's function is valid for frequencies near the resonances $\omega \approx \omega_\mu$ as long as the spectrum consists of non-overlapping resonance peaks, i.e., when the spacing between resonant modes exceeds the modal linewidths. First, let us rewrite the left-hand side of (A6) as an operator \mathcal{L}_ω acting on the field $\mathbf{E}(\mathbf{x}, \omega)$:

$$\mathcal{L}_\omega \mathbf{E}(\mathbf{x}, \omega) \equiv (\nabla \times \nabla \times - \omega^2 \hat{\varepsilon}_0(\omega, a_0)) \mathbf{E}(\mathbf{x}, \omega). \quad (\text{A24})$$

Next, we choose a complete set (see below) of eigenfunctions $\mathbf{E}_n(\mathbf{x}, \omega)$ and eigenvalues $\lambda_n(\omega)$ of the operator \mathcal{L}_ω :

$$\mathcal{L}_\omega \mathbf{E}_n(\mathbf{x}, \omega) = \lambda_n(\omega) \mathbf{E}_n(\mathbf{x}, \omega). \quad (\text{A25})$$

We define the inner product of two vector fields, $\mathbf{A}(\mathbf{x})$ and $\mathbf{B}(\mathbf{x})$, as $(A, B) \equiv \int dx \mathbf{A}(\mathbf{x}) \cdot \mathbf{B}(\mathbf{x})$. The operator \mathcal{L}_ω is complex symmetric under this inner product, i.e., $(A, \mathcal{L}_\omega B) = (\mathcal{L}_\omega A, B)$ [33, 65]. Therefore, we use unconjugated inner products throughout the derivation. In order to treat the set $\{\mathbf{E}_n\}$ as a discrete (countable) basis, a convenient theoretical trick is to place the system in a

box with absorbing boundary layers in which the absorption turns on more and more gradually. This procedure also gives the states \mathbf{E}_n finite norms (E_n, E_n). Because the operator is non-Hermitian, completeness of the basis can break down at an “exceptional point” [65, 87], but exceptional points are not generically present—they must be forced by careful tuning of parameters. Therefore, we assume completeness in this manuscript and will treat the influence of exceptional points (self-orthogonal modes) as a limiting case in a future paper, as discussed in Sec. VIII.

Let $G(\omega, \mathbf{x}, \mathbf{x}')$ be the Green’s function of the operator \mathcal{L}_ω , defined via $\mathcal{L}_\omega G(\omega, \mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}')$ [91]. Given the complete set of eigenfunctions and eigenvalues $\{\mathbf{E}_n, \lambda_n\}$, the Green’s function can be expressed as the sum [91]

$$G(\omega, \mathbf{x}, \mathbf{x}') = \sum_n \frac{\mathbf{E}_n(\mathbf{x}) \mathbf{E}_n^T(\mathbf{x}')}{\lambda_n(\omega) \cdot \int dx \mathbf{E}_n^2(\mathbf{x})}. \quad (\text{A26})$$

Each lasing mode is associated with an eigenvalue $\lambda_\mu(\omega)$ of \mathcal{L}_ω , which has a zero at a real frequency $\omega = \omega_\mu$. Consequently, $G(\omega, \mathbf{x}, \mathbf{x}')$ has a pole at ω_μ and at frequencies near ω_μ , it is dominated by a single term in the sum. Expanding $\lambda_\mu(\omega)$ around the pole $\lambda_\mu(\omega) \approx (\omega - \omega_\mu) \lambda'_\mu$ (where $\lambda'_\mu \equiv \frac{\partial \lambda}{\partial \omega}|_{\omega_\mu}$), we obtain

$$G_{\mu\mu}(\omega, \mathbf{x}, \mathbf{x}') \approx \frac{\mathbf{E}_\mu(\mathbf{x}) \mathbf{E}_\mu^T(\mathbf{x}')}{(\omega - \omega_\mu) \lambda'_\mu \cdot \int dx \mathbf{E}_\mu^2(\mathbf{x})}. \quad (\text{A27})$$

In order to evaluate λ'_μ , let us rewrite \mathcal{L}_ω as $\mathcal{L}_\omega \approx \mathcal{L}_{\omega_\mu} + V(\omega)$, where $\mathcal{L}_{\omega_\mu} \equiv \nabla \times \nabla \times - \omega_\mu^2 \hat{\varepsilon}_0(\omega_\mu)$ and $V(\omega) \equiv -[\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu (\omega - \omega_\mu)$. According to the Hellmann–Feynman theorem, the derivative of the eigenvalue $\lambda_\mu(\omega)$ with respect to ω is given by

$$\lambda'_\mu = \frac{\int dx \mathbf{E}_\mu^2(\mathbf{x}) [-\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu}{\int dx \mathbf{E}_\mu^2(\mathbf{x})}, \quad (\text{A28})$$

and substituting (A28) in (A27), we find that for frequencies near the resonances $\omega \approx \omega_\mu$, the Green’s function is approximately

$$G_{\mu\mu}(\mathbf{x}, \mathbf{x}', \omega) \approx \frac{\mathbf{E}_\mu(\mathbf{x}) \mathbf{E}_\mu^T(\mathbf{x}')}{(\omega_\mu - \omega) \int dx \mathbf{E}_\mu^2(\mathbf{x}) [\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu}. \quad (\text{A29})$$

5. The N-SALT TCMT equations

Having derived an expression for the Green’s function, the noise-driven field can be found by integrating the

Green’s function over the source terms $\hat{\mathbf{F}}_{\text{NL}}(\mathbf{x}', \omega)$ and $\hat{\mathbf{F}}_{\text{S}}(\mathbf{x}', \omega)$:

$$\begin{aligned} \hat{\mathbf{E}}_\mu(\mathbf{x}, \omega) &= \sum_\nu \omega_\mu^2 \int dx' G(\mathbf{x}, \mathbf{x}', \omega) \chi_\nu(\omega_\mu, a_0) \widehat{\Delta a}_\nu * \hat{\mathbf{E}}_\mu \\ &+ \int dx' G(\mathbf{x}, \mathbf{x}', \omega) \hat{\mathbf{F}}_{\text{S}}. \end{aligned} \quad (\text{A30})$$

In the first term on the right-hand side, we approximate $\omega \approx \omega_\mu$ because the correction term is $\mathcal{O}[(\omega - \omega_\mu) \cdot \widehat{\Delta a}]$, which is second order in the noise. Substituting the single-pole approximation (A29) in (A30) yields

$$\begin{aligned} \mathbf{E}_\mu \hat{a}_\mu &= \sum_\nu \frac{\omega_\mu^2}{(\omega - \omega_\mu)} \frac{\mathbf{E}_\mu \int \chi_\nu(\omega_\mu, a_0) \mathbf{E}_\mu^2}{\int dx \mathbf{E}_\mu^2(\mathbf{x}) [\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu} \widehat{\Delta a}_\nu * \hat{a}_\mu \\ &+ \frac{\mathbf{E}_\mu}{\omega - \omega_\mu} \frac{\int \hat{\mathbf{F}}_{\text{S}}(\mathbf{x}', \omega) \mathbf{E}_\mu}{\int dx \mathbf{E}_\mu^2(\mathbf{x}) [\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu}. \end{aligned} \quad (\text{A31})$$

Finally, multiplying both side by $\omega - \omega_\mu$ and taking the inverse Fourier transform, we arrive at the N-SALT TCMT equations, which govern the evolution of the modal amplitudes a_μ :

$$\begin{aligned} \dot{a}_\mu &= \sum_\nu \int dx c_{\mu\nu}(\mathbf{x}) \\ &\gamma(x) \int^t dt' e^{-\gamma(\mathbf{x})(t-t')} (|a_{\nu 0}|^2 - |a_\nu(t')|^2) a_\mu + f_\mu(t). \end{aligned} \quad (\text{A32})$$

The nonlinear coupling coefficient is

$$c_{\mu\nu}(\mathbf{x}) \equiv -i\omega_\mu^2 \frac{\chi_\nu(\omega_\mu, a_0) \mathbf{E}_\mu^2}{\int dx \mathbf{E}_\mu^2(\mathbf{x}) [\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu}, \quad (\text{A33})$$

and the Langevin force is

$$f_\mu(t) \equiv i \frac{\int [\mathbf{F}_{\text{S}}(\mathbf{x}', t) e^{-i\omega_\mu t}] \mathbf{E}_\mu}{\int dx \mathbf{E}_\mu^2(\mathbf{x}) [\omega^2 \hat{\varepsilon}_0(\omega)]'_\mu}. \quad (\text{A34})$$

APPENDIX B: LINEWIDTH OF THE MULTIMODE TIME-DELAYED MODEL

In this section, we calculate the laser linewidth for the multimode time-delayed model by generalizing the solution strategy of Sec. V in the text. We begin our analysis with the discretized time-delayed N-SALT TCMT equation (35) (repeated here for convenience):

$$\begin{aligned} \dot{a}_\mu &= \\ \sum_{\nu k} C_{\mu\nu}^k &\left[\gamma_k \int^t dt' e^{-\gamma_k(t-t')} (|a_\nu(t')|^2 - |a_{\nu 0}|^2) \right] a_\mu + f_\mu. \end{aligned} \quad (\text{B1})$$

Following the approach of Sec. V.A, we linearize (B1) by expanding the mode amplitudes a_μ around their steady-state values: $a_\mu = (a_{\mu 0} + \delta_\mu)e^{i\phi_\mu}$ (where $\delta_\mu \ll a_{\mu 0}$), and we omit the terms $\mathcal{O}(\delta_\mu^2)$. Then, we introduce additional variables ξ_μ^k

$$\xi_\mu^k = \gamma_k \int^t dt' e^{-\gamma_k(t-t')} \delta_\mu(t'), \quad (\text{B2})$$

where ξ_μ^k is the time-averaged amplitude deviation of mode $\mu = 1 \dots M$ from steady state at the spatial point $k = 1 \dots N$. Having introduced the auxiliary variables ξ_μ^k , the set of integro-differential equations (B1) turns into a linear system of ODEs, which we solve by applying several linear-algebraic transformations to obtain a compact expression for the covariance matrix, as described in detail below.

Introducing the vector $\Phi_\mu \equiv a_{\mu 0} \phi_\mu$, the linear system of ODEs is conveniently written as

$$\dot{\delta}_\mu = - \sum_{\nu k} (2a_{\mu 0} a_{\nu 0} \text{Re}[C_{\mu\nu}^k]) \xi_\nu^k + f_\mu^R, \quad (\text{B3})$$

$$\dot{\Phi}_\mu = - \sum_{\nu k} (2a_{\mu 0} a_{\nu 0} \text{Im}[C_{\mu\nu}^k]) \xi_\nu^k + f_\mu^I \quad (\text{B4})$$

$$\dot{\xi}_\mu^k = -\gamma_k \xi_\mu^k + \gamma_k \delta_\mu. \quad (\text{B5})$$

To simplify the notation further, we introduce the $M \times M$ matrices \mathbb{A}^k and \mathbb{B}^k ($k = 1 \dots N$), with entries

$$A_{\mu\nu}^k = 2a_{\mu 0} a_{\nu 0} \text{Re}[C_{\mu\nu}^k] \quad (\text{B6})$$

$$B_{\mu\nu}^k = 2a_{\mu 0} a_{\nu 0} \text{Im}[C_{\mu\nu}^k], \quad (\text{B7})$$

and we rearrange the set of equations (B3-B5) in a matrix form [compare with (24,25)]:

$$\frac{d}{dt} \boldsymbol{\delta} = - \sum_k \mathbb{A}_k \boldsymbol{\xi}^k + \mathbf{f}^R, \quad (\text{B8})$$

$$\frac{d}{dt} \boldsymbol{\Phi} = - \sum_k \mathbb{B}_k \boldsymbol{\xi}^k + \mathbf{f}^I, \quad (\text{B9})$$

$$\frac{d}{dt} \boldsymbol{\xi}^k = -\gamma_k \boldsymbol{\xi}^k + \gamma_k \boldsymbol{\delta}. \quad (\text{B10})$$

The autocorrelation matrix of the phase vector $\boldsymbol{\Phi}$, which we calculate in this section, is determined by the autocorrelation matrix of the Langevin force

$$\langle \mathbf{f}(t) \mathbf{f}^{*T}(t') \rangle = \mathbb{R} \delta(t - t'). \quad (\text{B11})$$

In order to compute $\langle \boldsymbol{\Phi} \boldsymbol{\Phi}^T \rangle$, we solve (B9) by straightforward integration. We find that the phase covariance matrix is a sum of a “pure” phase-diffusion term, proportional to $\frac{\mathbb{R}}{2}$, and an amplitude-phase coupling term, proportional to \mathcal{J} :

$$\langle \boldsymbol{\Phi}(t) \boldsymbol{\Phi}^T(0) \rangle = \left(\frac{\mathbb{R}}{2} + \mathcal{J} \right) |t|, \quad (\text{B12})$$

where we have introduced the shorthand notation

$$\mathcal{J} \equiv \frac{1}{|t|} \sum_{kl} \mathbb{B}_k \iint dt' ds' \langle \xi_k(t') \xi_l^T(s') \rangle \mathbb{B}_l^T \quad (\text{B13})$$

for the second term, which is responsible for the generalized α factor.

In the remainder of this section, we calculate \mathcal{J} . First, we solve the set of ODEs for $\boldsymbol{\xi}^k$ and $\boldsymbol{\delta}$ (B8,B10), and then we substitute the solution for $\boldsymbol{\delta}$ into (B13) and evaluate the integrals. To this end, we begin by rewriting the equations for $\boldsymbol{\xi}^k$ and $\boldsymbol{\delta}$ more compactly. We define the $[(N+1) \cdot M] \times 1$ vectors \mathbf{x} and \mathbf{F} and the $[(N+1) \cdot M] \times [(N+1) \cdot M]$ matrix \mathbb{K} :

$$\mathbf{x} = \begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\xi}^1 \\ \vdots \\ \boldsymbol{\xi}^N \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} \mathbf{f}^R \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}$$

$$\mathbb{K} = \begin{pmatrix} 0 & \mathbb{A}_1 & \mathbb{A}_2 & \dots & \mathbb{A}_N \\ \mathbb{A}_1 & -\mathbb{A}_1 & 0 & \dots & 0 \\ \mathbb{A}_2 & 0 & -\mathbb{A}_2 & & 0 \\ \vdots & \vdots & 0 & \ddots & 0 \\ \mathbb{A}_N & 0 & \dots & 0 & -\mathbb{A}_N \end{pmatrix}, \quad (\text{B14})$$

where \mathbb{A}_k are block-diagonal $M \times M$ matrices with γ_k on the diagonal entries, and the zeros in the definition of \mathbb{K} are block $M \times M$ zero matrices. Using these definitions, the equations for $\boldsymbol{\xi}^k$ and $\boldsymbol{\delta}$ (B8,B10) can be conveniently written as

$$\frac{d}{dt} \mathbf{x} = -\mathbb{K} \mathbf{x} + \mathbf{F}. \quad (\text{B15})$$

The solution of (B15) is

$$x_m(t) = \int^t dt' \sum_\rho \left[e^{-\mathbb{K}(t-t')} \right]_{m\rho} F_\rho(t') \quad (\text{B16})$$

and, in particular, the solution for ξ_μ^k is

$$\xi_\mu^k = \int^t dt' \sum_{s=1}^M \left[e^{-\mathbb{K}(t-t')} \right]_{Mk+\mu,s} f_s^R(t'). \quad (\text{B17})$$

For ease of notation, let us denote the $(k+1)^{\text{st}}$ $M \times M$ block in the first column of the matrix $e^{-\mathbb{K}(t-t')}$ by the shorthand notation $[e^{-\mathbb{K}(t-t')}]_{k+1,1}$, so that $\xi_k = \int_0^t dt' [e^{-\mathbb{K}(t-t')}]_{k+1,1} \mathbf{f}^R(t')$. Substituting the expression for ξ_k into \mathcal{J} and using the autocorrelation function of the Langevin force (B11), we obtain

$$\mathcal{J} = \frac{1}{|t|} \sum_{kl} \mathbb{B}_k \iiint dt' dt'' ds' \left[e^{-\mathbb{K}(t'-t'')} \right]_{k+1,1} \times \frac{\mathbb{R}}{2} \left[e^{-\mathbb{K}^T(s'-t'')} \right]_{\ell+1,1} \mathbb{B}_\ell^T. \quad (\text{B18})$$

We proceed (not shown) by diagonalizing the matrix \mathbb{K} and evaluating the integrals in (B18). (The intermediate steps depend on the eigenvalues of \mathbb{K} and the matrix of eigenvectors, but the final result can be expressed in terms of the matrix inverse \mathbb{K}^{-1}). In the long-time limit, we keep the leading order term (which grows linearly in time) and we obtain

$$\mathcal{J} = \left(\sum_k \mathbb{B}_k [\mathbb{K}]_{k+1,1}^{-1} \right) \frac{\mathbb{R}}{2} \left(\sum_\ell (\mathbb{B}_\ell [\mathbb{K}]_{\ell+1,1}^{-1})^T \right). \quad (\text{B19})$$

In order to complete the derivation of the linewidth formula, we use the following identity:

$$[\mathbb{K}]_{k+1,1}^{-1} = \left(\sum_j \mathbb{A}_j \right)^{-1}, \quad (\text{B20})$$

which we prove below. Noting that $A_{\mu\nu} = \sum_k A_{\mu\nu}^k$ and $B_{\mu\nu} = \sum_k B_{\mu\nu}^k$ and using the identity (B20), we find that (B19) reduces to

$$\mathcal{J} = \mathbb{B} \mathbb{A}^{-1} \frac{\mathbb{R}}{2} (\mathbb{B} \mathbb{A}^{-1})^T, \quad (\text{B21})$$

which completes the derivation of the linewidth formula in the most general time-delayed model. In particular, and somewhat remarkably, the γ terms completely cancel in the computation of the first column of the matrix inverse, and drop out of the final result.

Proof of the identity (B20): We use Schur complement [92] for the lower-left corner of a matrix inverse:

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{pmatrix}^{-1} = \begin{pmatrix} * & * \\ -\mathbb{D}^{-1}\mathbb{C}(\mathbb{A} - \mathbb{B}\mathbb{D}^{-1}\mathbb{C})^{-1} & * \end{pmatrix},$$

(\mathbb{A} and \mathbb{D} need to be square matrices). Decomposing the matrix \mathbb{K} into the blocks

$$\mathbb{A} = \begin{pmatrix} 0 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} \mathbb{A}_1 & \mathbb{A}_2 & \dots & \mathbb{A}_N \end{pmatrix},$$

$$\mathbb{C} = \begin{pmatrix} \mathbb{A}_1 \\ \mathbb{A}_2 \\ \vdots \\ \mathbb{A}_N \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} -\mathbb{A}_1 & 0 & \dots & 0 \\ 0 & -\mathbb{A}_2 & & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & -\mathbb{A}_N \end{pmatrix},$$

we can calculate the lower-left corner of $[\mathbb{K}]_{k+1,1}^{-1}$:

$$-\mathbb{D}^{-1}\mathbb{C}(\mathbb{A} - \mathbb{B}\mathbb{D}^{-1}\mathbb{C})^{-1} = (\mathbb{D}^{-1}\mathbb{C})[\mathbb{B}(\mathbb{D}^{-1}\mathbb{C})]^{-1} =$$

$$-\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \left[\begin{pmatrix} \mathbb{A}_1 & \mathbb{A}_2 & \dots & \mathbb{A}_N \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right]^{-1} =$$

$$-\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \left[\sum_I \mathbb{A}_I \right]^{-1}.$$

Therefore, we obtain $[\mathbb{K}]_{k+1,1}^{-1} = \left(\sum_j \mathbb{A}_j \right)^{-1}$.

APPENDIX C: LEMMA FROM SEC. VI.C

$$[\mathbb{B} \mathbb{A}^{-1}]_{ij} \propto \frac{a_{i0}}{a_{j0}}$$

In Sec. VI.B, we present a multimode Schawlow–Townes relation, which states that the linewidths are proportional to a sum of inverse output powers of all the other modes. This result arises from a lemma which we prove here. We use the standard matrix-inverse formula [93]

$$\mathbb{A}^{-1} = \frac{1}{\det \mathbb{A}} \text{adj} \mathbb{A}, \quad (\text{C1})$$

where the adjugate matrix is defined as

$$\text{adj} \mathbb{A} = ((-1)^{i+j} M_{ij})^T. \quad (\text{C2})$$

\mathbb{M} is the cofactor matrix, i.e. the matrix whose (i, j) entry is the determinant of the (i, j) minor of \mathbb{A} (which is the matrix obtained from \mathbb{A} by deleting the i 'th row and the j 'th column). From the definition of \mathbb{A} (i.e., $A_{ij} \equiv \text{Re}[C_{ij}]a_{i0}a_{j0}$), it follows that

$$M_{ij} = \left(\prod_{k \neq i, j} a_{k0}^2 \right) a_{i0}a_{j0} Q_{ij}. \quad (\text{C3})$$

where Q_{ij} (and later Q) denote constants that may depend on i and j , but are independent of the modal amplitudes. Note also that

$$\det \mathbb{A} = \prod_k a_{k0}^2 \cdot Q. \quad (\text{C4})$$

Using (C1-C4), we obtain

$$A_{ij}^{-1} = \frac{1}{a_{i0}a_{j0}} \cdot Q_{ij}. \quad (\text{C5})$$

Therefore, one can easily see that the lemma follows, since $[\mathbb{B} \mathbb{A}^{-1}]_{ij} \propto \sum_k a_{i0}a_{k0} \cdot \frac{1}{a_{k0}a_{j0}} \propto \frac{a_{i0}}{a_{j0}}$.

APPENDIX D: COMPARISON WITH THE SCATTERING-MATRIX LINewidth FORMULA

In a recent scattering-matrix based linewidth theory [14], Pillay *et al.* obtain a formula for the linewidth of a one-dimensional laser system, expressed in terms of integrals over the modes which solve the nonlinear SALT

equations. In this appendix, we prove that their formula (which applies to 1d systems) is equivalent to our linewidth formula (3) (except that their formula gives a spatially averaged incomplete-inversion factor and omits the α factor).

In the scattering-matrix approach, the lasing modes are described as purely outgoing wave functions ψ_0 , which satisfy the nonlinear SALT equation

$$\nabla \times \nabla \times \psi_0(x) - \omega_0^2 \varepsilon(x, \omega_0) \psi_0(x) = 0, \quad (\text{D1})$$

and can be expressed as a superposition of outgoing channel modes \mathbf{u}_μ outside of the laser region

$$\psi_0(x) = \sum_k b_k \mathbf{u}_k(x, \omega_0) \quad \text{for } x \notin C. \quad (\text{D2})$$

C denotes the scattering region (i.e., $\varepsilon = 1$ for $x \notin C$). Note that $\psi_0(x)$ is precisely the same the mode $\mathbf{E}_0(\mathbf{x})$ (which was used in Sec. II) inside the cavity region.] The outgoing mode-amplitudes \mathbf{b} are normalized to the value of ψ_0 at the cavity boundary ($x = L$)

$$\mathbf{b}^T \mathbf{b} = \psi_0^2(L). \quad (\text{D3})$$

The apparent difference between our formula and the linewidth formula in [14] is that the integral term in the denominator of our linewidth formula(3) is replaced by a sum of two terms in the scattering-matrix approach

$$\int_{\text{all space}} dx \left[\varepsilon \omega_0 + \frac{\omega_0^2}{2} \frac{d\varepsilon}{d\omega_0} \right] \psi_0^2 \longrightarrow \frac{i\mathbf{b}^T \mathbf{b}}{2} + \int_C dx \left[\varepsilon \omega_0 + \frac{\omega_0^2}{2} \frac{d\varepsilon}{d\omega_0} \right] \psi_0^2. \quad (\text{D4})$$

In order for the two formulas to agree, we need to show that

$$\omega_0 \int_L^\infty dx \psi_0^2(x) = \frac{i\mathbf{b}^T \mathbf{b}}{2} \quad (\text{D5})$$

(where we have used the fact that $\varepsilon = 1$ outside the cavity region). We show that the latter condition (D5) holds for any solution ψ_0 of (D1) which satisfies outgoing boundary conditions. One way to impose outgoing boundary conditions is to invoke the limiting-absorption principle (i.e., add loss to eliminate incoming waves from infinity and take the limit of infinitesimal absorption at the end of the calculation [94, 95]). Formally, we define the integral on the left-hand side of (D5) as

$$\int_L^\infty dx \psi_0^2(x) \equiv \lim_{s \rightarrow 0^+} \int_L^\infty dx e^{-sx} \psi_0^2(x). \quad (\text{D6})$$

By substituting $\psi_0(x) = e^{ik_0 x}$ into (D6) and taking the limit of $s \rightarrow 0^+$, we obtain $\int_L^\infty dx \psi_0^2(x) = \frac{i}{2k} e^{2ik_0 L} = \frac{i}{2k} \mathbf{b}^T \mathbf{b}$, and since $\omega_0 = ck_0$ this finishes the proof of (D3) (with the units convention of $c = 1$).

APPENDIX E: ZERO-POINT FLUCTUATION CANCELLATION

The hyperbolic cotangent factor in the FDT (19) arises as a sum of a Bose–Einstein distribution and a $1/2$ factor stemming from quantum zero-point (ZP) fluctuations [70, 71], and this is why it does not vanish in the limit of zero temperature ($\beta \rightarrow \infty$). However, it turns out that contribution of this ZP term cancels in the linewidth formula, as was shown by Henry and Kazarinov [72] from a quantum-operator viewpoint, and it is convenient to explicitly subtract the ZP term from the hyperbolic cotangent as in (3) and (21). Here, we provide a purely classical explanation for why this cancellation occurs, and why it is important to perform the explicit subtraction in order to eliminate a subtlety arising from the definition of outgoing boundary conditions.

The FDT has a hyperbolic cotangent factor, and when we apply the FDT to find the $\langle f_\mu f_\nu^* \rangle$ correlation function in Sec IV, the same hyperbolic cotangent factor arises in the R integral, appearing in the form

$$\begin{aligned} & \int dx |\mathbf{E}_\mu|^2 \text{Im } \varepsilon_0(\omega_\mu) \cdot \frac{1}{2} \coth \left(\frac{\hbar \omega_\mu \beta}{2} \right) \\ &= \int dx |\mathbf{E}_\mu|^2 \text{Im } \varepsilon_0(\omega_\mu) \left[\left(\frac{1}{2} \coth \frac{\hbar \omega_\mu \beta}{2} - \frac{1}{2} \right) + \frac{1}{2} \right] \end{aligned} \quad (\text{E1})$$

for a lasing mode μ , where we have trivially added and subtracted the ZP $1/2$ factor from \coth . Now, we wish to analyze the final $1/2$ term, which is the integral $\frac{1}{2} \int dx |\mathbf{E}_\mu|^2 \text{Im } \varepsilon_0(\omega_\mu)$. Before we treat outgoing boundary conditions, let us consider the simpler case of a laser surrounded by an explicit absorbing medium, as in [72]. (This is also the situation in more recent computational models, for which one uses a finite spatial domain surrounded by absorbing layers [21].) For any steady-state lasing mode (real ω_μ), the net gain + loss is zero, but $\frac{1}{2} \int dx |\mathbf{E}_\mu|^2 \text{Im } \varepsilon_0(\omega_\mu)$ is proportional to the net power absorbed or gained by the electric field [96] and hence this integral is *zero*. Therefore, in such a case, whether or not we include the $1/2$ factor is irrelevant, because the $\pm 1/2$ terms integrate to zero.

However, a subtlety arises in this integral in the common case where the laser is surrounded by an infinite zero-temperature ($\beta = \infty$) lossless medium with outgoing radiation boundary conditions. Outgoing boundary conditions can be defined mathematically by the limiting absorption principle [94, 95]: one takes the lossless medium to be the limit of a lossy medium as the losses go to zero from above, which can be expressed by writing ε as $\varepsilon + i0^+$. Just as in appendix D, the correct approach is to take the lossless limit *after* solving the problem, i.e. the 0^+ limit is taken *outside* of the integral. Before we take this limit, it makes no difference whether the $1/2$ factor is included, just as above: it integrates to zero. However, after we take the lossless limit, there is no explicit absorbing region ($\text{Im } \varepsilon > 0$) in the

integral (the absorption has been “moved to infinity” in some sense), so if we perform the coth integral without subtracting $1/2$ then we would obtain an incorrect contribution from the ZP fluctuations in the gain medium (which should have been canceled). Instead, if we integrate against $1/2\coth - 1/2$, the result is correct without requiring any explicit contribution from the absorbing boundary conditions.

Note that if the laser is surrounded by an infinite lossless medium at a *positive* temperature, then there is a nonzero contribution of incoming thermal radiation to

the linewidth [2, 3]. This can be included in one of two ways. In practice, we typically solve the SALT equations in a finite computational box with an explicit absorbing region, in which case no modification to our linewidth formula is required: one simply assigns the ambient temperature to the absorbing region. If, on the other hand, the outgoing boundary conditions are imposed in some other way (e.g. semi-analytically as in earlier SALT work [19, 20, 43, 97]), then an explicit source term must be added to account for incoming thermal radiation, as in previous works [98].

-
- [1] M. Sargent, M. O. Scully, and W. E. J. Lamb, *Laser Physics* (Westview Press, 1974).
 - [2] H. Haken, *Laser Theory* (Springer-Verlag, 1984).
 - [3] H. Haken, *Laser Light Dynamics* (North-Holland, 1985).
 - [4] O. Svelto, *Principles of Lasers* (Springer, 1976).
 - [5] P. W. Milonni and J. H. Eberly, *Laser Physics* (Wiley, 2010).
 - [6] J. P. Gordon, H. J. Zeiger, and C. H. Townes, *Phys. Rev.* **99**, 1264 (1955).
 - [7] A. L. Schawlow and C. H. Townes, *Phys. Rev.* **112**, 1940 (1958).
 - [8] K. Petermann, *IEEE J. Quant. Elect.* **15**, 566 (1979).
 - [9] M. Lax, *Physics of Quantum Electronics*, edited by P. Kelley, M. Lax, and P. Tannenwald (McGraw-Hill, New York, 1966).
 - [10] M. P. van Exter, S. J. M. Kuppens, and J. P. Woerdman, *Phys. Rev. A* **51**, 809 (1995).
 - [11] C. H. Henry, *IEEE J. Quant. Elect.* **18**, 259 (1982).
 - [12] M. Osinski and J. Buus, *IEEE J. Quant. Elect.* **23**, 9 (1987).
 - [13] Y. D. Chong and A. D. Stone, *Phys. Rev. Lett.* **109**, 063902 (2012).
 - [14] J. C. Pillay, N. Yuki, A. D. Stone, and Y. D. Chong, *Phys. Rev. A* **89**, 033840 (2014).
 - [15] L. He, Şahin Kaya Özdemir, and L. Yang, *Laser Photonics Rev.* **7**, 60 (2013).
 - [16] O. Painter, R. Lee, A. Scherer, A. Yariv, J. D. O’Brien, P. Dapkus, and I. Kim, *Science* **284**, 1819 (1999).
 - [17] M. Loncar, T. Yoshie, A. Scherer, P. Gogna, and Y. Qiu, *Appl. Phys. Lett.* **74**, 7 (1999).
 - [18] H. G. Park, S. Kim, S. Kwon, Y. Ju, J. Yang, J. Baek, S. Kim, and Y. Lee, *Science* **305**, 1444 (2004).
 - [19] H. E. Türeci, A. D. Stone, and B. Collier, *Phys. Rev. A* **74**, 043822 (2006).
 - [20] L. Ge, Y. D. Chong, and A. D. Stone, *Phys. Rev. A* **82**, 063824 (2010).
 - [21] S. Esterhazy, D. Liu, M. Liertzer, A. Cerjan, L. Ge, K. Makris, A. D. Stone, J. Melenk, S. G. Johnson, and S. Rotter, *Phys. Rev. A* **90**, 023816 (2014).
 - [22] W. E. Lamb, *Phys. Rev.* **134**, A1429 (1964).
 - [23] H. B. Callen and T. A. Welton, *Phys. Rev.* **81**, 34 (1951).
 - [24] S. M. Rytov, *Principles of Statistical Radiophysics II: Correlation Theory of Random Processes* (Springer-Verlag, 1989).
 - [25] I. E. Dzyaloshinskii, E. M. Lifshitz, and L. P. Pitaevskii, *Sov. Phys. Usp.* **4**, 153 (1961).
 - [26] E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics: Part 2* (Pergamon-Oxford, 1980).
 - [27] W. Eckhardt, *Opt. Comm.* **41**, 305 (1982).
 - [28] M. Lax, *Phys. Rev.* **160**, 290 (1967).
 - [29] C. H. Henry, *J. Lightwave Tech.* **LT-4**, 288 (1986).
 - [30] H. A. Haus, *Waves and Fields in Optoelectronics* (Englewood Cliffs, NJ:Prentice-Hall, 1984).
 - [31] H. A. Haus and W. Huang, *Proc. of the IEEE* **79**, 1505 (1991).
 - [32] W. Suh, Z. Wang, and S. Fan, *IEEE J. Quant. Elect.* **40**, 1511 (2004).
 - [33] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals, Molding the Flow of Light* (Princeton University Press, 2008).
 - [34] A. Rodriguez, M. Soljačić, J. D. Joannopoulos, and S. G. Johnson, *Opt. Exp.* **15**, 7303 (2007).
 - [35] F. T. Arecchi, G. L. Lippi, G. P. Puccioni, and J. R. Tredicce, *Opt. Comm.* **51**, 308 (1984).
 - [36] G. L. Oppo, A. Politi, G. L. Lippi, and F. T. Arecchi, *Phys. Rev. A* **34**, 4000 (1986).
 - [37] L. A. Lugiato, P. Mandel, and L. M. Narducci, *Phys. Rev. A* **29**, 1438 (1984).
 - [38] M. P. van Exter, W. A. Hamel, J. P. Woerdman, and B. R. P. Zeijlmans, *IEEE J. Quant. Elect.* **28**, 1470 (1992).
 - [39] W. Elsasser, *IEEE J. Quant. Elect.* **21**, 687 (1985).
 - [40] R. Hui, S. Benedetto, and I. Montrosset, *IEEE J. Quant. Elect.* **29**, 1488 (1993).
 - [41] A. Cerjan, A. Pick, Y. Chong, S. G. Johnson, and A. D. Stone, *arXiv:1505.01884*.
 - [42] D. S. Wiersma, *Nat. Phys.* **4**, 359 (2008).
 - [43] H. E. Türeci, L. Ge, S. Rotter, and A. D. Stone, *Science* **320**, 643 (2008).
 - [44] G. H. Duan, P. Gallion, and G. Debarge, *IEEE J. Quant. Elect.* **26**, 32 (1990).
 - [45] A. E. Siegman, *Lasers* (University Science Books, 1986) p. 462.
 - [46] M. Patra, *On Quantum Optics in Random Media*, Ph.D. thesis, University of Leiden (2005).
 - [47] S. J. M. Kuppens, *Opt. Comm.* **107**, 249 (1994).
 - [48] S. J. M. Kuppens, M. A. van Eijkelenborg, M. P. v. E. C. A. Schrama, and J. P. Woerdman, *IEEE J. Quant. Elect.* **32**, 383 (1996).
 - [49] A. E. Siegman, *Phys. Rev. A* **39**, 1264 (1989).
 - [50] S. J. M. Kuppens, M. P. van Exter, and J. P. Woerdman, *Phys. Rev. Lett.* **72**, 3815 (1994).
 - [51] S. J. M. Kuppens, M. P. van Exter, M. van Duin, and J. P. Woerdman, *IEEE J. Quant. Elect.* **31**, 1237 (1995).
 - [52] K. Vahala, L. C. Chiu, S. Margalit, and A. Yariv, *Appl.*

- Phys. Rev. **42**, 631 (1983).
- [53] L. D. Westbrook and M. J. Adams, IEE Proc. J. Optoelectron. **134**, 209 (1987).
 - [54] R. J. Lang and A. Yariv, IEEE J. Quant. Elect. **22**, 436 (1986).
 - [55] J. Arnaud, Opt. Quant. Elect. **18**, 335 (1986).
 - [56] H. Schomerus, K. M. Frahm, M. Patra, and C. W. J. Beenakker, Physica A **278**, 469 (2000).
 - [57] H. Schomerus, Phys. Rev. A **79**, 061801(R) (2009).
 - [58] H. E. Türeci, A. D. Stone, L. Ge, S. Rotter, and R. J. Tandy, Nonlinearity **22**, C1 (2008).
 - [59] A. E. Siegman, Phys. Rev. A **39**, 1253 (1989).
 - [60] M. O. Scully, G. Süssmann, and C. Benkert, Phys. Rev. Lett. **60**, 1014 (1988).
 - [61] M. Scully, M. Zubairy, and K. Wódkiewicz, Opt. Comm. **65**, 440 (1988).
 - [62] C. Benkert, M. O. Scully, A. A. Rangwala, and W. Schleich, Phys. Rev. A **42**, 1487 (1990).
 - [63] C. Benkert, M. O. Scully, and G. Süssmann, Phys. Rev. A **41**, 6119 (1990).
 - [64] M. I. Kolobov, L. Davidovich, E. Giacobino, and C. Fabre, Phys. Rev. A **47**, 1431 (1993).
 - [65] N. Moiseyev, *Non-Hermitian Quantum Mechanics* (Cambridge University Press, 2011) pp. 174–183.
 - [66] A. E. Siegman, *Frontiers of Laser Physics and Quantum Optics* (Springer, 2000) pp. 31–37.
 - [67] J. R. Jeffers, N. Imoto, and R. Loudon, Phys. Rev. A **47**, 3346 (1993).
 - [68] R. Matloob, R. Loudon, M. Artoni, S. M. Barnett, and J. Jeffers, Phys. Rev. A **55**, 1623 (1997).
 - [69] M. Patra and C. W. J. Beenakker, Phys. Rev. A **60**, 4059 (1999).
 - [70] L. D. Landau and E. M. Lifshitz, *Statistical Physics, 3rd edition, Part I* (Elsevier Ltd., 1980) p. 386.
 - [71] C. Kittel and H. Kroemer, *Thermal Physics* (W. H. Freeman and Company, 1980).
 - [72] C. H. Henry and R. F. Kazarinov, Rev. Mod. Phys. **68**, 801 (1996).
 - [73] A. Cerjan, Y. Chong, L. Ge, and A. D. Stone, Opt. Exp. **20**, 474 (2012).
 - [74] A. Cerjan, Y. D. Chong, and A. D. Stone, Opt. Exp. **23**, 6455 (2015).
 - [75] B. Dana, L. Lobachinsky, and A. Bahabad, Opt. Comm. **324** (2014).
 - [76] M. P. van Exter, W. A. Hamel, J. P. Woerdman, and B. J. P. Zeijlmans, IEEE J. Quant. Elect. **28**, 1470 (1992).
 - [77] A. Raman and S. Fan, Phys. Rev. B **83**, 205131 (2011).
 - [78] W. Feller, Bull. Amer. Math. Soc. **51**, 800 (1945).
 - [79] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., Vol. 1 (New York: Wiley, 1968).
 - [80] See Eq. 45 in Ref. [29].
 - [81] A. Demir, A. Mehrotra, and J. Roychowdhury, IEEE Trans. Circuits Syst. I, Fundam. Theory Appl. **47**, 655 (2000).
 - [82] See Eq. A4 in Ref. [20].
 - [83] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes, the art of scientific computing* (Cambridge University Press, 2007) p. 907.
 - [84] A. Pick et al. In preparation.
 - [85] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics, 2nd Edition* (Wiley, 2007).
 - [86] S. Y. Lin, E. Chow, S. G. Johnson, and J. D. Joannopoulos, Opt. Lett. **26**, 1903 (2001).
 - [87] W. D. Heiss, J. Phys. A: Math. Theor. **45**, 444016 (2010).
 - [88] M. Liertz, L. Ge, A. Cerjan, A. D. Stone, H. E. Türeci, and S. Rotter, Phys. Rev. Lett. **108**, 173901 (2012).
 - [89] M. Brandstetter, M. Liertz, C. Deutsch, P. Klang, J. J. Schöberl, H. E. Türeci, G. Strasser, K. Unterrainer, and S. Rotter, Nat. Comm. **5**, 4034 (2014).
 - [90] S. Y. Lee, J. W. Ryu, J. B. Shim, S. B. Lee, S. W. Kim, and K. An, Phys. Rev. A **78**, 015805 (2008).
 - [91] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Elsevier Academic Press, 2006) pp. 662–663.
 - [92] B. A. Bultheel and M. V. Bare, *Linear Algebra, Rational Approximation and Orthogonal Polynomials* (North-Holland, 1997) p. 87.
 - [93] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists* (Elsevier Academic Press, 2006) pp. 184–185.
 - [94] J. R. Schulenberger and C. H. Wilcox, Arch. Ration. Mech. Anal. **41**, 46 (1971).
 - [95] A. Taflove, A. Oskooi, and S. G. Johnson, *Advances in FDTD Computational Electrodynamics: Photonics and Nanotechnology* (Artech House, 2013) p. 76.
 - [96] J. D. Jackson, *Classical Electrodynamics Third Edition* (John Wiley and Sons, 1999) pp. 262–264.
 - [97] H. E. Türeci, A. D. Stone, and L. Ge, Phys. Rev. A **76**, 013813 (2007).
 - [98] J. Andreasen, H. Cao, A. Taflove, P. Kumar, and C. Cao, Phys. Rev. A **77**, 023810 (2008).